CRYSTAL STRUCTURES OF NITROALKANE OXIDASE: INSIGHTS INTO THE STRUCTURAL BASIS FOR SUBSTRATE SPECIFICITY AND THE CATALYTIC MECHANISM

A Thesis
Presented to
The Academic Faculty

by

Akanksha Nagpal

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in Chemistry

Georgia Institute of Technology
August 2005

CRYSTAL STRUCTURES OF NITROALKANE OXIDASE: INSIGHTS INTO THE STRUCTURAL BASIS FOR SUBSTRATE SPECIFICITY AND THE CATALYTIC MECHANISM

Approved by:

Dr. Allen M. Orville, Advisor
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Loren D. Williams
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Giovanni Gadda
Department of Chemistry
Georgia State University

Dr. Donald F. Doyle
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Dale E. Edmondson
Department of Chemistry
Emory University

Date Approved: 11th July 2005

Dedicated

to my loving and caring husband, Sudeep, and our newborn, Aayush, for their endless love, support and patience

My inspiration: My Mom, who reared me to have strong determination

and taught me to always strive for the best

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere appreciation and gratitude to my advisor, Dr. Allen M. Orville, for his invaluable guidance and patience on all the facets of this thesis work. His technical expertise, his boundless excitement for research and creativity, and his attention to detail were constant sources of inspiration and motivation. He always encouraged me to come up with creative and independent ideas.

I am thankful to our collaborators Dr. Fitzpatrick and Dr. Valley at Texas A&M University for providing the purified protein samples for crystallization. I want to especially thank Dr. Fitzpatrick for providing insightful discussion on NAO structures. I also want to thank my committee members, Dr. Loren D. Williams, Dr. Donald F. Doyle, Dr. Dale E. Edmondson and Dr. Giovanni Gadda, for providing valuable suggestions during this work and for serving on my committee.

Most of the data sets presented in this thesis were collected at NSLS and APS synchrotron sources. I am grateful to the members of beamline X26-C especially Dr. Annie Heroux, Dr. Robert Sweet, and Dr. Dieter Schneider for helping me with data collection during my trips to NSLS. I am thankful to APS for granting me GUP time for data collection on the long cell edged crystals of NAO. I also received a great deal of support from beamline scientists Dr. Ning Lei and Dr. Keith Brister at BIOCARS, APS for data collection. I would also like to acknowledge Dr. Reinhard Pahl at BIOCARS, who introduced me to the idea of *crystal annealing*, which worked like a charm for long

cell edged NAO crystals. I am thankful to IMCA-CAT beamline director Dr. Lisa Keefe and beamline scientist Irina Koshelev for helping me with MAD data collection on different crystals of NAO. I would also like to thank *CCP4, CNS, PyMoL* discussion board members, who shared their experiences and advice on topics related to crystal structure solution.

I am grateful to members of my research group, George Lountos and Linda Manning, for making my stay at Georgia Tech a great leaning experience. I remember innumerable synchrotron trips with George and his help and support during those trips. I also appreciate George for reading my thesis chapters and providing me with helpful suggestions. I would like to thank Wijaya Martanto for help with thesis formatting. It was also great to have Tatsuya Mahegashi, Derrick Watkins, Dr. Seiji Komeda, Lopak, Dr. Tinoush Moulaei, Srividya (Dr. Loren William's lab) as neighbors.

I would like to thank Dr. Loren D. Williams, Dr. Cameron J. Tyson, Dr. George McKelvy, and Dr.Toby Block for giving me lighter teaching load when I needed it the most. I also want to thank Tatsuya Mahegashi for helping me with Webct TA assignment, which saved me a lot of time for thesis writing.

Finally, I would like to thank my family and parents for their constant love and support. My mother is the greatest mom in the world. She has always been a constant source of inspiration to me and has always stood by my decisions, especially the one of

coming to US. I miss the meals cooked by my father when I used to work for long hours in the lab after Aayush was born. Although miles away, my parents have never let a moment go without praying for my well being and my success. My brother Gaurav and sister-in-law Naina have always encouraged and supported me in my endeavors. My niece, Gaurika (a.k.a Nikku), has always charmed me with her cute activities (....*Aashuu Bua, Kaiche ho*....). I would also like to thank my sister-in-law Neelu and brother-in-law Rakesh for wishing me success and luck. Last but not the least, my loving and caring husband, Sudeep, always kept me going during all the lows in life. I am the luckiest to have him as my husband. Aayush, my wonderful son, made the biggest sacrifice for this thesis. He didn't see me for two months (when he was just four months old) and stayed with his father in California just so I could concentrate on my Ph.D. and wrap up my thesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Nitrochemicals are widely distributed throughout the environment from natural and human sources. They are present in rocket fuels, wastes from chemical industry, in the exhaust from internal combustion engines and tobacco smoke. They are also used as explosives, herbicides, pesticides, biocides and drugs. In addition, 3-nitro-tyrosine and other nitrated protein residues are important post-translational markers for oxidative damage, septic shock, Alzheimer's disease as well as many cardiovascular, neurodegenerative and malignant conditions. Nitrated proteins, amino acids and metabolites often possess antibiotic activity and, consequently, can mitigate bacterial and/or fungal infection. Because of the wide presence of the nitrocompounds as toxins, potential nitrogen/carbon sources or metabolic intermediates, different organisms have evolved to produce enzymes that can biodegrade them. Nitroalkane oxidase (NAO) produced by *Fusarium oxysporum*, is one such enzyme.

NAO catalyzes the oxidation of the neutral nitroalkanes to the corresponding aldehydes or ketones with production of H_2O_2 and nitrite. The enzyme is overexpressed when cultures of the fungus are grown on nitroethane as a sole source of nitrogen. The reasonable physiological roles for NAO include: a) that it protects the fungus from nitroalkanes produced by plants, b) to detoxify antibiotics produced by the competing soil microbes such as *Pseudomonas fluorescens*, and/or c) as a strategy to scavenge nitrogen from a variety of nitrochemicals. NAO is a unique flavoenzyme, which belongs to acyl-CoA dehydrogenase (ACAD) superfamily but does not bind or turnover acyl-CoA substrates. The enzyme is also different from its ACAD homologs in following a

carbanion reaction mechanism. Interestingly, the resting state of the enzyme selectively reacts with neutral nitroalkanes but after a few steps into the catalytic cycle, the enzyme reacts with negatively charged nitroalkane species to form a stable covalent N5-FAD adduct. The adduct has been isolated and characterized as N5-(2-nitrobutyl)-1,5-dihydro-FAD in case of nitroethane substrate.

Most of the biochemical studies on NAO were performed by our collaborators (Dr. Fitzpatrick, Dr. Valley and Dr. Gadda) but the novel characteristics of the enzyme prompted us to initiate structural studies of the enzyme. X-ray diffraction methods along with the spectroscopic techniques were used to solve the three dimensional molecular structure of resting state (E^{ox}), inhibited state (EI) and substrate trapped reaction intermediate (ES*) of NAO. The crystal structures of EI and ES* of NAO were solved to 2.07 Å and 2.2 Å resolutions respectively.

The homotetrameric structure of ES* was solved by MAD phasing with 52 Se-Met sites in an orthorhombic space group. The electron density for the N5-(2-nitrobutyl)-1,5-dihydro-FAD covalent intermediate was clearly resolved. The structure of ES* was used to solve the crystal structure of EI, a complex with weak competitive inhibitor, spermine. The crystals of EI grew with a *c* axis of 485 Å and 1½ holoenzymes in the asymmetric unit, which made high resolution data collection very challenging. The structures of ES* and EI reveal a hydrophobic channel that extends from the exterior of the protein to the *re*-face of the FAD which terminates at the N5-FAD position. Both the structures also show that Asp402 is located at the end of the substrate access channel. Moreover, the

residue is in the correct position to serve as the active site base, where it is proposed to abstract the α-proton from neutral nitroalkane substrates. The protein fold for NAO and various members of the ACAD superfamily overlay with root mean square deviations between 1.7 to 3.1 Å. The homologous region typically spans more than 325 residues and includes Glu376, which is the active site base in the prototypical member of the ACAD family. However, NAO and the ACADs exhibit differences in substrate access channel, hydrogen bonding interactions between the respective active site base, substrate molecules and the FAD. These aspects differentiate NAO from its homologs and, therefore are proposed to result in the unique reaction mechanism of NAO.

The crystal structures of NAO have given us structural insights into the reaction mechanism of this new member of ACAD superfamily. At the same time, these crystal structures also form the basis for directed evolution studies. It is likely that NAO and other ACAD superfamily members might have diverged from a common ancestor to perform different chemistries but retained an overall similar protein fold especially since *Fusarium oxysporum* has high mutation rate and a large variety of transposons. The molecular structure of NAO in conjunction with protein engineering techniques can also be used for designing biocatalysts for different substrates. Another interesting application of NAO can be in the field of biosensors. The enzyme has a versatile redox center, which can be modulated by the active site environment. The modified FAD can then serve as on-off or gradient switch in various bioelectronic applications.

1 INTRODUCTION

1.1 Nitrocompounds in Biosphere

Nitrochemicals from natural and anthropomorphic sources are distributed throughout the environment. Organic nitro-compounds are often difficult to metabolize, mutagenic or toxic and consequently occur only rarely from natural sources. However, when they are biosynthesized they are often used as a chemical defense against pathogens (Clark *et al.*, 1974; Hanawa *et al.*, 2000). For example, 3-nitropropionic acid (3NPA), which is produced by some plants, is extremely toxic and inhibits succinate dehydrogenase and fumarase with very high affinity. Plants like *Hippocrepis comosa* produce 3NPA in extremely large quantities in their leaves, which serves as powerful chemical defense mechanism (Hipkin *et al.*, 1999). The biocidal activity of 3NPA resides in part, by inhibition of the citric acid cycle of potential pathogens (Porter & Bright, 1980; Hassel & Sonnewald, 1995). Recently, 3NPA has also been shown to induce oxidative stress in striatal and brain cortex synaptosomes in rats (Tunez *et al.*, 2005).

Another class of organic nitro-compounds is the antibiotics secreted by various prokaryotes and eukaryotes. Broad range antibiotics like chloramphenicol bind to 50S ribosomal unit and inhibit peptidyl transferase activity in bacteria (Marconi *et al.*, 1990). Azomycin, a nitroheterocyclic antibiotic produced by *Streptomyces* and some strains of *Pseudomonas* is the basis for the most popular synthetic drug metronidazole (Osato *et al.*, 1955; Shoji *et al.*, 1989). Apart from toxins and antibiotics, the toxicity of many nitrochemicals originates from their metabolic transformation. The reduction of organic

nitrochemicals by one, two or four electrons produces nitro-radical anions, nitroso and hydroxylamino intermediates, respectively (Nathan & Shiloh, 2000). These types of reactive nitrogen intermediates are strong mutagens similar to nitrous acid and hydroxylamine, which readily modify DNA bases. For example, nitroreduction of dehydrochloramphenicol (DH-CAP), a bacterial metabolite of chloramphenicol, by human bone marrow cells produces toxins that are the underlying cause of genotoxicity (Isildar *et al.*, 1988).

In addition, nitrated protein residues, such as 3-nitro-tyrosine (3-NO_2-Tyr) are now important post-translational markers for oxidative damage, septic shock, Alzheimer's disease as well as many cardiovascular, neurodegenerative and malignant conditions (Aulak *et al.*, 2004; Irie *et al.*, 2003; Schopfer *et al.*, 2003; Turko & Murad, 2002). Explosive compounds like nitroglycerin (TNG) and 2,4,6-trinitrotoluene (TNT) do not have any known naturally occurring counterparts; therefore, they are often considered xenobiotic (Spain, 1995; Spain *et al.*, 2000; Rosser *et al.*, 2001).

1.2 Biotransformation of Nitrochemicals

Recent studies have established that many nitrochemicals are catabolized by prokaryotic and eukaryotic organisms as sources of nitrogen and/or carbon (Spain, 1995; Spain *et al.*, 2000; Rosser *et al.*, 2001; French *et al.*, 1998; Trott *et al.*, 2003; Somerville *et al.*, 1995; Ebert *et al.*, 2001). Indeed, several prokaryotic, FMN-dependent enzymes have received recent attention because they reductively eliminate nitrite from xenobiotic

explosive compounds, which can serve as a sole sources of nitrogen (Spain, 1995; French *et al.*, 1998; Johnson *et al.*, 2002; Binks *et al.*, 1996; Khan *et al.*, 2002; Meah *et al.*, 2001; Nivinskas *et al.*, 2000; Williams *et al.*, 2004). Enzymes like xenobiotic reductase A isolated from *Pseudomonas putida* II-B (XenA) and B from *P. fluorescens* I-C (XenB) catalyze NADPH-dependent reductive nitrite elimination from nitroglycerin. These strains not only tolerate concentrations of nitroglycerin that are toxic to nearly all other microbes, but they can also obtain all their nitrogen for growth from the xenobiotic compound (Blehert *et al.*, 1997; Blehert *et al.*, 1999). Thus it appears that the reductive enzymes enable some microorganisms to exploit nitrochemicals introduced into the environment by recent human activities.

In contrast, several FAD-dependent, eukaryotic enzymes are known to catalyze oxidative nitrite eliminations from naturally occurring nitroaliphatic compounds (Dhawale, M. R. & Hornemann, U., 1979; Kido *et al.*, 1976; Kido *et al.*, 1976; Kido *et al.*, 1978; Kido & Soda, 1978; Gadda *et al.*, 2000; Gadda & Fitzpatrick, 1999). The enzymes in this class appear to have evolved in parallel with nitrochemical biosynthesis pathways. Typically, the organisms that produce nitroalkane toxins such as 3-nitropropionate, also produce enzymes that transform the nitroaliphatic internally, and thus protect the host from the toxic effects of the agent (Porter & Bright, 1987; Hipkin *et al.*, 1999). In turn, some of the organisms targeted by the toxins also evolved strategies to detoxify various nitrocompounds (Lee *et al.*, 2003).

1.3 Nitrite Elimination by Flavoenzyme Oxidases

Flavoenzymes are a ubiquitous and diverse class of biological redox catalysts (Muller, 1991). The redox activity centers of these enzymes are derivatives of riboflavin (vitamin B2), which consists of an isoalloxazine ring system substituted with methyl groups at C7 and C8 and a ribityl group at N10 (Stankovich, 1991). The flavin redox center is chemically versatile and flavoenzymes can act, for example, as oxidases, monooxygenases, electron transferases, dehydrogenases, disulphide reductases etc. (Massey, 2000). There are several examples of oxidative flavoenzymes that transform nitroalkanes. These include, 2-nitropropane dioxygenase (2-NPD) from *Hansenula mrakii* (Kido, 1984) or *Neurospora crassa* (Gorlatova, 1998; Francis *et al.*, 2005), nitroalkane dioxygenase from *Streptomyces ansochromogenes* (Zhang & Tan, 2002), 3-nitropropionate oxidase (3-NPO) from either *Penicillium atrovenetum* or *Hippocrepis comosa* (Porter & Bright, 1987; Hipkin *et al.*, 1999), nitroalkane oxidase from *Fusarium oxysporum* (Kido *et al.*, 1978b) (NAO) or glucose oxidase from *Aspergillus niger* (Swoboda & Massey, 1965), and D-amino acid oxidase from mammals (Alston *et al.*, 1983). Most of the nitrite eliminating enzymes only react with the anionic form of the nitro substrate except for 2-NPD from *Neurospora crassa*, which reacts with the neutral and as well as the anionic nitroalkanes (Francis *et al.*, 2005). However, NAO is unique since it only catalyzes the neutral nitroalkanes (Gadda *et al.*, 1997). The reactions catalyzed with nitro-containing substrates by glucose oxidase and D-amino acid oxidase are not physiological, since they do not appear to provide a benefit to the organisms (Kido *et al.*, 1978a). In contrast, the reactions catalyzed by each of the other oxidases do provide obvious advantages, such as the ability to obtain all nitrogen required for growth

from a variety of nitrochemicals or to protect the microbes against host defense mechanism.

NAO from *Fusarium oxysporum* has been very well characterized in the recent years (Fitzpatrick *et al.*, 2005). The enzyme is produced in large quantities when cultures of the fungus obtain all the nitrogen required for growth from nitroethane (Kido *et al.*, 1978b). Reasonable physiological roles for the enzyme include: a) that it protects the fungus from nitroalkanes produced by plants (Alston *et al.*, 1977), b) to detoxify antibiotics produced by the competing soil microbes such as *Pseudomonas fluorescens* (Lee *et al.*, 2003), and/or c) as a strategy to scavenge nitrogen from a variety of nitrochemicals. The molecular weight of the 439 residue subunit is 47,955 and analytical ultracentrifugation suggests that the enzyme has a quaternary structure comprised of a homotetramer/homodimer mixture (Gadda & Fitzpatrick, 1998). The enzyme may be located in the peroxisome, since it has amino acid sequence homology in the C-terminal region with other enzymes targeted for the peroxisome (Neuberger *et al.*, 2003). Furthermore, amino acid sequence comparisons also indicate that NAO is homologous to the acyl-CoA dehydrogenase (ACAD) superfamily of flavoproteins, with similarities ranging between 7 - 46 % (Fitzpatrick *et al.*, 2005; Ghisla & Thorpe, 2004; Kim & Miura, 2004; Daubner *et al.*, 2002). NAO oxidizes a variety of neutral nitro-substrates, but does not catalyze acyl-CoA substrates, amino or hydroxy acids (Gadda & Fitzpatrick, 1998). NAO has a marked preference for unbranched, primary nitroalkanes with four to six carbon atoms (Gadda & Fitzpatrick, 1999). For example, the k_{cat}/K_M value for NAO increases with linear nitroaliphatic substrates with an increasing number of carbon atoms

and maximizes with nitrohexane. Indeed, for short primary nitroalkanes, each additional CH_2 moiety increases the k_{cat}/K_m value by 2.5 kcal/mol, and that 1.7 kcal/mol appears to be used to increase the binding affinity and the remainder for catalysis (Gadda & Fitzpatrick, 1999).

Mechanistic studies of the reaction cycle of NAO, in particular with nitroethane as the substrate, have established many of the details of the catalytic cycle (Fitzpatrick *et al.*, 2005). The reductive half-reaction is initiated with the abstraction of the substrate alpha proton by Asp402 (Figure 1, step 1). Formation of the nitroethane anion is rate-limiting and 2-$^2H_2$-nitroethane exhibits a k_{cat}/K_m deuterium isotope effect of 7.5 - 7.9 (Gadda *et al.*, 2000d; Valley *et al.*, 2005). Thus, a carbanion is definitively established as an intermediate in the catalytic cycle (Gadda & Fitzpatrick, 2000e), another unique feature of the NAO reaction among flavoenzymes. Moreover, at 25 °C, NAO catalyzes the ionization of nitroethane a billion-fold more effectively than does acetate in an analogous, non-enzymatic reaction (Valley & Fitzpatrick, 2004). The increase in rate is primarily due to a large decrease in the enthalpy of activation, in accord with the concept that enzymes generally catalyze reactions by decreasing $\Delta H^{\ddagger}$ as described by Wolfenden (Wolfenden *et al.*, 1999). The resulting nitroethyl-anion adds to the flavin N5 to form the first covalent intermediate (Figure 1, step 2). The loss of nitrite and subsequent addition of hydroxide leads to aldehyde product formation and reduced enzyme (Figure1, steps 3-5). In the absence of excess nitroethyl anion, the reduced enzyme is rapidly oxidized by O_2 to yield H_2O_2 as illustrated by steps 6-7. Alternatively, a stable adduct can be formed by addition of a second nitroethane anion to the electrophilic imine intermediate of the

original substrate molecule (step 8, ES*). Mass spectrometric analysis of the flavin species released from ES* at low pH is consistent with a N5-(2-nitrobutyl)-1,5-dihydro-FAD adduct (Gadda *et al.*, 1997). The decay of ES* *in vitro* via step 9 is slow at room temperature, pH 8.5, and is O_2 independent. However, it does yield oxidized enzyme and a similar *in vivo* reaction is also possible. Consequently, we describe the ES* complex as a true reaction cycle intermediate. The remarkable stability of ES* at 4 °C and neutral pH, and its lack of reactivity with O_2, enabled us to crystallize the ES* form of the enzyme.

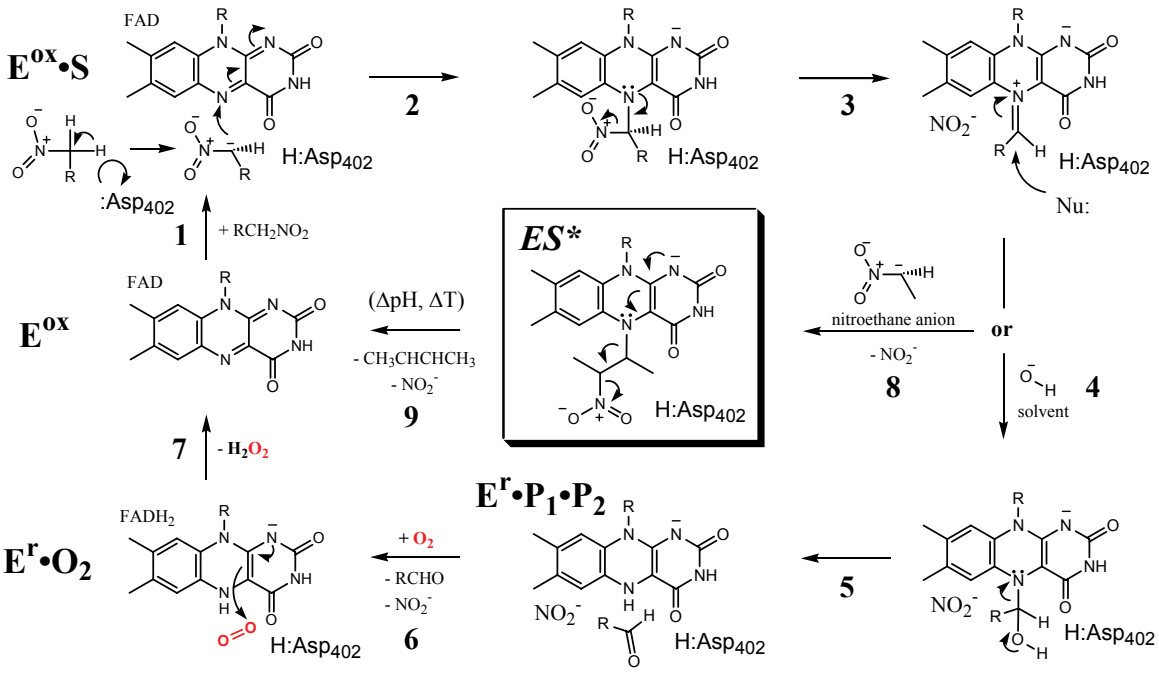


Figure 1.1 The proposed NAO reaction mechanism with nitroethane as substrate (Gadda & Fitzpatrick, 2000e). The imine reaction intermediate after step 3 can react with hydroxide to release the aldehyde followed by oxidation by O_2 (steps 4-7). Alternatively, the intermediate can react with a nitroalkane anion (step 8) to generate the ES* complex. It converts to the active enzyme in a pH and/or temperature-dependent, but O_2 independent manner (step 9) (Nagpal *et al.*, 2005).

We solved the structure with Se-Met enriched protein and MAD phasing techniques (Hendrickson & Ogata, 1997) using the 52 Se-Met sites of the homotetramer. To our knowledge, this is the first crystal structure of a flavoprotein trapped during turnover of authentic substrates as a covalent adduct at N5-FAD. We also report the structure of oxidized NAO, in which we observe a spermine molecule bound in some of the active sites. Together, these structures provide insights into: a) substrate preference and access to the FAD, b) the reaction mechanism, and c) the features that differentiate NAO from structural, but not functional homologs in the ACAD family.

2 BACKGROUND & SIGNIFICANCE

2.1 Diversity and Utility of Nitrochemicals

Nitrocompounds are widely used as drugs, herbicides, pesticides and explosives (Figure 2.1) (Ahlner *et al.*, 1991; Spain, 1995). They are also present in the exhaust of internal combustion engines and tobacco smoke. Consequently, they have been distributed throughout the environment from man-made sources in vast excess over naturally occurring examples. Naturally occurring nitro-compounds include simple compounds like nitropropionic acid or more complex examples such as chloramphenicol and azomycin. These compounds were first characterized approximately 50 years ago and are among the first recognized examples of antibiotics. They also have inspired the production of numerous synthetic nitrochemicals. For example, metronidazole is a leading drug for the treatment of human ulcers caused by *Helicobacter pylor*i. The pathogen contains the *rdxA* gene, which encodes for an NADPH-dependent, oxygen insensitive nitroreductase (Jeong *et al.*, 2000). The enzyme reduces the nitro group of metronidazole (-415 mV reduction potential) (Mendz & Megraud, 2002) to yield the cytotoxic derivatives that kill the pathogen (Sisson *et al.*, 2000; Sisson *et al.*, 2002; Trinh & Reysset, 1998; Samuelson, 1999). Unfortunately, there are increasing numbers of *H. pylori* strains with altered reactivity or expression of the *rdxA* gene product that are metronidazole resistant (Freeman *et al.*, 1997; Goodwin *et al.*, 1998; Mendz & Megraud, 2002). Explosive compounds such as nitroglycerin (TNG), 2,4,6-trinitrotoluene (TNT), RDX and CL-20 do not have any known naturally occurring counterparts; therefore, they are often considered xenobiotic (Spain *et al.*, 2000; Rosser *et al.*, 2001; Bottaro, 1996).

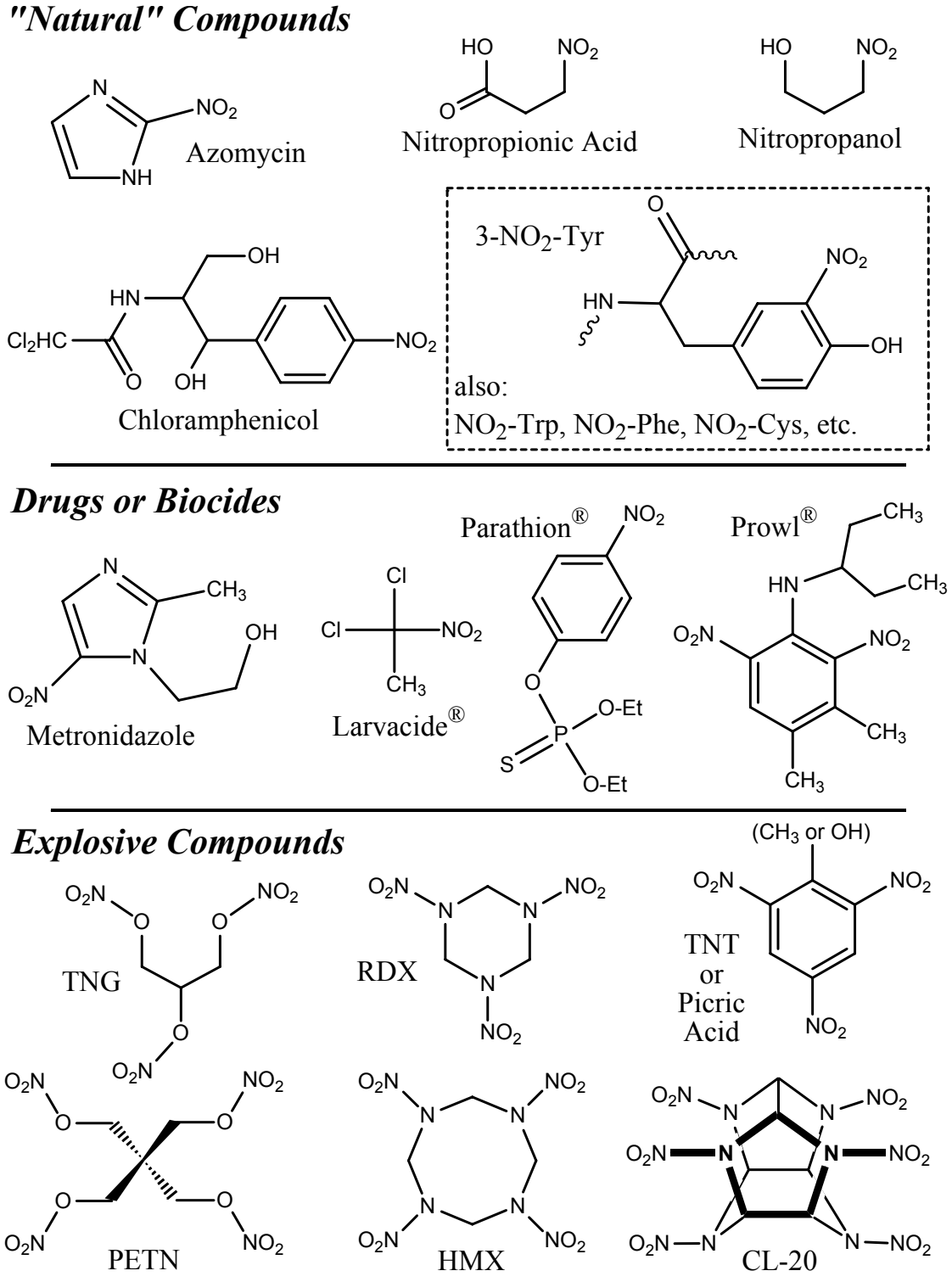


Figure 2.1 Chemical structures and names of some nitrochemicals dispersed throughout the environment by natural and man-made sources.

Nitroglycerin was first synthesized by Sobrero in 1846 and clearly predates the more powerful explosives RDX (~1920) and CL-20 (>1975). Many explosives are toxic at concentrations well below their solubility limit (Homma-Takeda *et al.*, 2002). For example, Prowl ® (Figure 2.1) and related nitroaromatic explosives and pesticides are proton transporting ionophores and thus uncouple oxidative phosphorylation. The presence of these compounds in the soil and ground water is a major risk factor at some EPA superfund sites. Nevertheless, their agricultural and military uses greatly extend their distribution throughout the environment with commensurate associated risks to human health.

2.2 Biodegradation of Nitrochemicals

The reductive biotransformation of nitro groups can yield reactive nitroso (R-NO) and hydroxylamino (R-NHOH) intermediates (Figure 2.2), which are capable of modifying DNA (Nathan & Shiloh, 2000). Reduction of a nitrochemical by six electrons is necessary to yield the amino derivative, which can be easily assimilated in proteins and nucleic acids. Thus nitrochemicals are often difficult to metabolize, mutagenic or toxic. For example, ethylnitrosourea, a potential product of reduction of some nitramines, is a strong alkylating agent that causes genetic transversions. Nevertheless, studies of microbes isolated from environments containing natural or xenobiotic nitrochemicals have established that most of these nitrochemicals are catabolized.

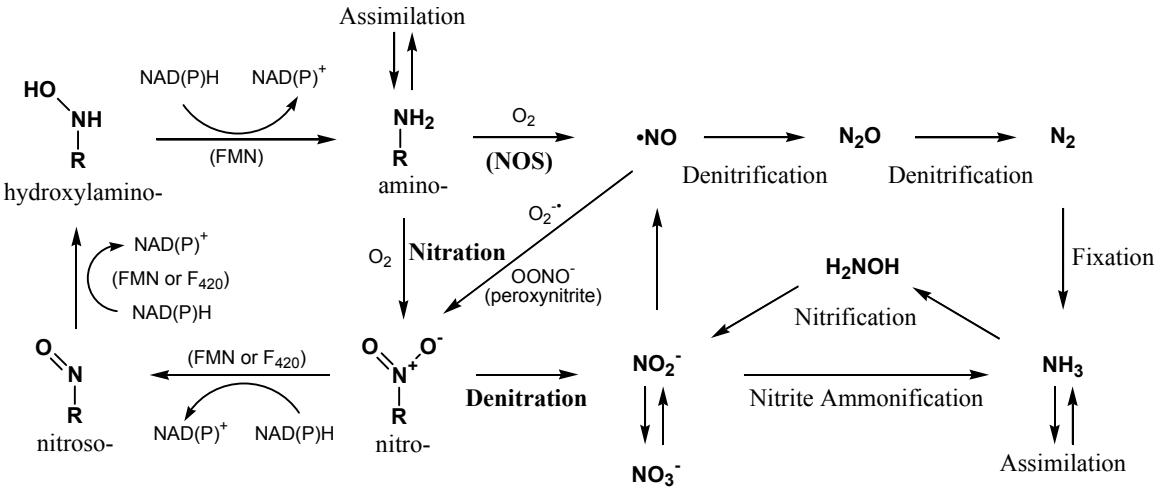


Figure 2.2 The nitrogen cycles. Enzymatic reductions of organic nitrochemicals (left portion) yield toxic nitroso- and hydroxylamino- intermediates, but amino compounds are assimilated. Denitration reactions yield inorganic nitrite that must be transformed into ammonia for assimilation (right portion). Formation of organic nitrochemicals occurs through at least two routes: via peroxynitrite (from O_2^- • and •NO) or by direct oxidation of an amino precursor.

Nitroesters (R-O-NO_2), nitroaromatics (Ar-NO_2), nitroalkanes (R-CH_2-NO_2) and nitramines (R_2-N-NO_2) each serve as sources of nitrogen, carbon and/or energy to some microbes (Spain, 1995; Rosser *et al.*, 2001; Spain *et al.*, 2000; Trott *et al.*, 2003; Bhushan *et al.*, 2003; Johnson *et al.*, 2002; Williams & Bruce, 2000; Coleman *et al.*, 2002). Reductive transformation of R-O-NO_2 typically yields an alcohol and eliminates nitrite (Spain *et al.*, 2000; Williams & Bruce, 2000, 2002; Williams *et al.*, 2001; Williams *et al.*, 2004; Binks *et al.*, 1996; Blehert *et al.*, 1997). Interestingly, the explosive compound, nitroglycerin is also currently one of the 100 most frequently prescribed drugs for the treatment of angina pectoris (Ahlner *et al.*, 1991). Mammalian transformation of nitroglycerin by vascular smooth muscle cells is known to yield 1,2-dinitroglycerol, nitrite, and eventually nitric oxide (•NO). Chen *et al.* demonstrated in 2002 that a novel

activity of mammalian mitochondrial aldehyde dehydrogenase (mt-ALDH) catalyzes the regiospecific, reductive denitration of nitroglycerin to yield nitrite and 1,2-dinitroglycerin (Chen *et al.*, 2002). In contrast, reduction of Ar-NO_2 tends to yield nitroso and hydroxylamino derivatives, but does not liberate nitrite (Spain, 1995; Rosser *et al.*, 2001; Johnson *et al.*, 2002; Pak *et al.*, 2000; Spiess *et al.*, 1998; Somerville *et al.*, 1995; French *et al.*, 1998). Oxidation of Ar-NO_2 by some flavoenzymes eliminates nitrite commensurate with hydroxylation (Suzuki *et al.*, 1991; Spain, 1995; Johnson *et al.*, 2002). Finally, oxidative nitrite elimination from nitroaliphatics occurs with incorporation of either solvent or molecular oxygen into the aldehyde or ketone products (Dhawale, M. R., and & Hornemann, U., 1979; Gadda *et al.*, 1997; Kido *et al.*, 1976a; Kido *et al.*, 1978b). Remarkably, *Rhodococcus erythropolis* HL PM-1 has very recently been shown to encode a complete catabolic pathway for picric acid and several enzymes in the pathway use coenzyme F420 (Figure 2.3), a 5-deaza-flavin analog (Heiss *et al.*, 2003; Rieger *et al.*, 2002; Heiss *et al.*, 2002; Ebert *et al.*, 2001). In addition, CL-20 has been recently shown to be transformed by the NADH-dependent, FAD containing enzyme, salicylate 1-monooxygenase from *Pseudomonas sp.* strain ATCC 29352 (Bhushan *et al.*, 2004). The physiological role of salicylate 1-monooxygenase is to biotransform salicylate to catechol but the enzyme also transforms CL-20, *o*-nitrophenol and benzoates at a lower reaction rate (Suzuki *et al.*, 1991). Therefore, although natural sources of nitrochemicals may appear scarce and nitrochemicals introduced by human activities are often considered toxic, they can be efficiently exploited by microbial communities. They may even be rapidly turned over by a heretofore uncharacterized "metabolic economy". The assimilation of priority pollutants found in some of EPA

superfund sites is eagerly anticipated. However, the catabolism of drugs by human pathogens is also occurring, which is not as desirable a scenario.

Organic nitrocompounds can be produced from precursor amines by two biological pathways, collectively called nitration. In the first pathway, the amino precursor of chloramphenicol, for example, is oxidized directly to the final nitro-containing antibiotic by an oxidase with an uncharacterized reaction mechanism (van Pee, 1996). The second pathway is responsible for the production of nitro-substituted amino acids such as 3-NO_2-tyrosine. This represents a potentially ubiquitous class of naturally occurring nitrochemicals (Nathan & Shiloh, 2000; MacMicking *et al.*, 1997; Bogdan, 2001). It is now recognized that 3-NO_2-Tyr and other nitrated protein residues are important post-translational markers for oxidative damage, septic shock, Alzheimer's disease as well as many cardiovascular, neurodegenerative and malignant conditions (Aulak *et al.*, 2004; Aulak *et al.*, 2001; Irie *et al.*, 2003; Kamisaki *et al.*, 1998; Schopfer *et al.*, 2003; Turko *et al.*, 2003). For example, mammalian phagocytes actively engaged in combating a microbial infection produce nitric oxide (generated by inducible nitric oxide synthase) and superoxide (O_2^- •, generated by NADPH phagocyte oxidase) (Nathan & Shiloh, 2000; Xu & Liu, 1998). Nitric oxide is normally derived from arginine by nitric oxide synthase and has many physiological functions, including vasodilatation (Nathan & Shiloh, 2000; Bogdan, 2001; MacMicking *et al.*, 1997; Xu & Liu, 1998). The •NO and O2 -• react to produce peroxynitrite (OONO -), which is proposed to generate 3-NO_2-Tyr and other nitrated amino acids in proteins (Aulak *et al.*, 2004; Aulak *et al.*, 2001; Turko *et al.*, 2003; Miyagi *et al.*, 2002). Nitrated proteins, amino acids and

metabolites may possess antibiotic activity and consequently mitigate bacterial and/or fungal infection (Turko & Murad, 2002; Kamisaki *et al.*, 1998). Recent proteomic studies suggest that the presence of 3-NO_2-Tyr in proteins is dynamically linked to the physiological status of the cell, tissue and organ (Aulak *et al.*, 2004; Turko *et al.*, 2003; Shishehbor & Hazen, 2004; Kuhn *et al.*, 2004; Gow *et al.*, 2004; Butterfield, 2004; Aulak *et al.*, 2004). More than 40 proteins are rapidly nitrated and denitrated as conditions change (Koeck *et al.*, 2004; Duncan, 2003; Ye *et al.*, 1996; van der Vliet *et al.*, 1996; Crow & Ischiropoulos, 1996; Aulak *et al.*, 2004). They are found in the cytosol and the mitochondria and the level of nitration is dependent upon the physiological status of the cell or tissue type. However, very little is known about the nitration and denitration mechanisms, the processes are likely to be analogous to the ubiquitous phosphorylation cycles of Ser, Thr and Tyr residues and metabolites in most life forms.

2.3 Flavoproteins are Ubiquitous in Biochemistry

As suggested above, many of the biological transformations of organic nitrochemicals are either catalyzed by, or otherwise involve flavoproteins. FAD and FMN are the most common organic cofactors found in enzymes (Massey, 2000; Fraaije & Mattevi, 2000; Ghisla & Massey, 1989). Consequently, they facilitate very diverse catalytic activities (Palfey & Massey, 1998). The isoalloxazine ring system in flavoenzymes is ideally suited for oxidative or reductive reactions involving one or two electron processes, as well as reactivity with molecular oxygen (Figure 2.3). Thus flavoproteins couple two-electron metabolic processes to one-electron transfer processes of respiration.

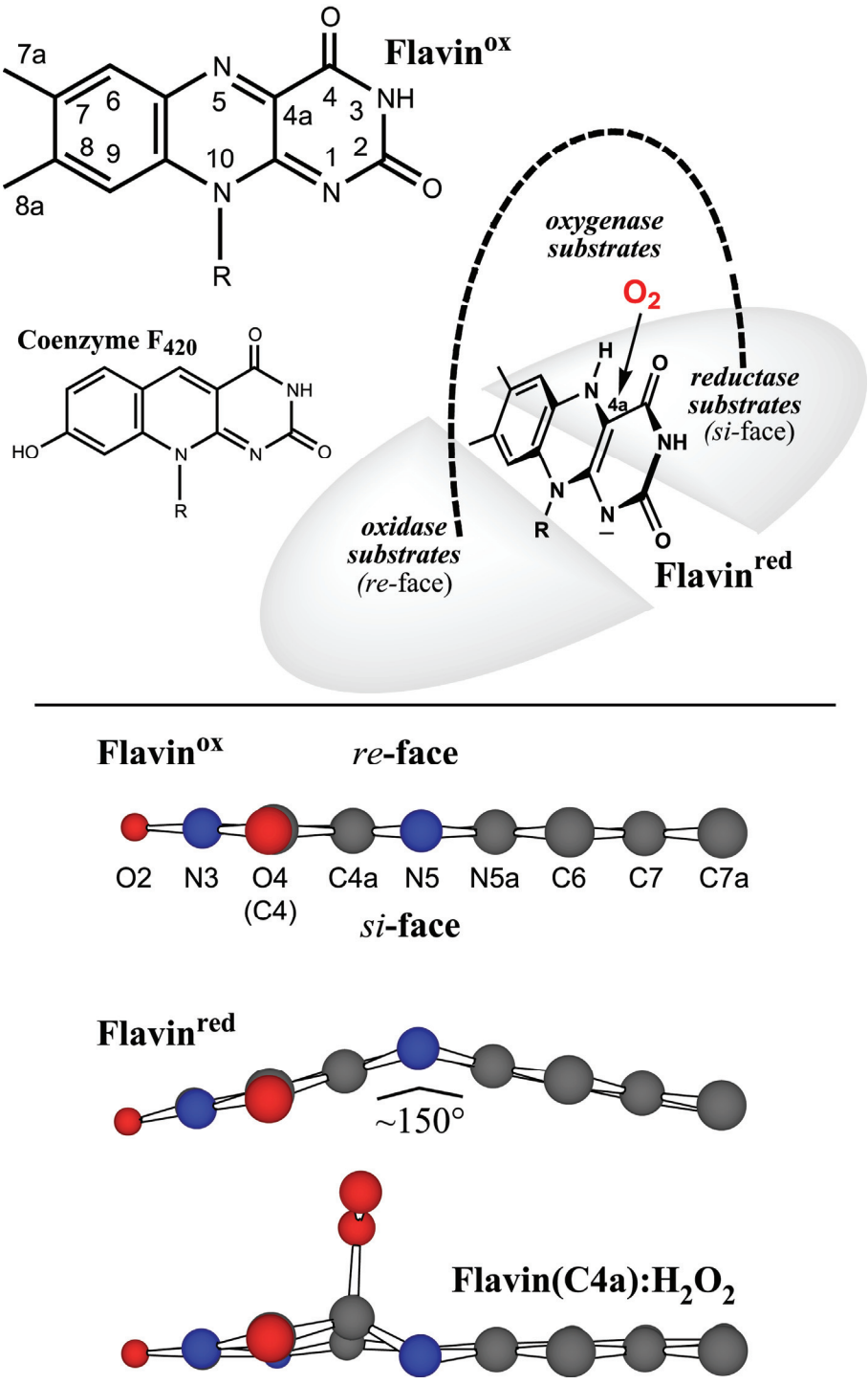


Figure 2.3 A structure-based hypothesis for reactivity in flavoenzymes. Depending upon the redox state, the isoalloxazine ring either shows a planar or bent conformation. The reaction of the reduced flavin with O_2 is likely to proceed via C4a-hydroperoxide intermediate formation (Bruice, 1980).

They are also involved in bioluminescence, programmed cell death, detoxification, DNA repair and photosynthetic processes. Moreover, the isoalloxazine ring facilitates catalysis in many types of enzymes because it can act as an electrophile, a nucleophile or form a covalent intermediate at either N5 or C4a. Not surprisingly then, FAD and FMN are among the most common organic cofactors found in biochemistry.

A myriad of fundamentally similar interactions influence each reaction coordinate and yet, yield remarkable catalytic diversity. Moreover, selective pressure and evolution have enhanced "desirable" reaction(s), while simultaneously decreasing deleterious reaction(s). Although more than 550 flavoprotein structures have been deposited in the PDB, much debate remains as to how an active site environment influences a given reaction coordinate. However, far less is known about how different reactions can occur with similar substrates and flavin cofactors. The study of the different enzymes that transform nitrochemicals, but employ different reaction mechanisms, will provide insights into these reaction processes.

The reaction(s) of a particular flavoprotein derive from a complex combination of interactions. These include the flavin cofactor itself, protein active site residues, substrate molecule(s) and inclusion or exclusion of solvent molecules and/or ions. A few well-studied flavoenzymes support proposals for several of these features. For example, structural analysis of numerous dehydrogenases, oxidases and reductases reveal that substrates often bind parallel to the isoalloxazine ring, but often on opposite faces (Figure 2.3) (Fraaije & Mattevi, 2000). The dehydrogenases and oxidases frequently bind

substrates adjacent to the *re*-face, whereas reductases often bind substrates parallel to the *si*-face. In contrast, *p*-hydroxybenzoate 3-hydroxylase and many other oxygenases, bind substrates near the top edge of the reduced flavin (Palfey *et al.*, 1999; Schreuder *et al.*, 1994; Gatti *et al.*, 1994; Schreuder *et al.*, 1992; Schreuder *et al.*, 1989; Schreuder *et al.*, 1988; Wierenga *et al.*, 1979). Thus the three-dimensional binding relationship between the substrate(s) and the cofactor profoundly influences the reaction. The reduction potential of the flavin is thought to be influenced by the presence of either positive or negative charge near the redox active N5 - N1 pair. For example, a positive charge near N1 appears to increase the reduction potential and may also stabilize the anionic state at N1 in the reduced cofactor (Ortiz-Maldonado *et al.*, 2004; Moran *et al.*, 1997; Wu *et al.*, 2003; Stankovich *et al.*, 1999; Johnson & Stankovich, 1993; Byron *et al.*, 1990; Fink *et al.*, 1986; Stankovich & Fox, 1984). Hydrogen bonding to N5 is thought to polarize the **L**owest **U**noccupied **M**olecular **O**rbital (LUMO) and (**H**ighest **O**ccupied **M**olecular **O**rbital (HOMO) of the N5 atom of oxidized flavins. Indeed, theoretical studies suggest that a hydrogen bond donor approximately in the plane of the isoalloxazine ring will orient the LUMO approximately perpendicular to the plane (Ghisla & Thorpe, 2004; Bach & Dmitrenko, 2004; Satoh *et al.*, 2003; Dmitrenko *et al.*, 2003; Bach *et al.*, 2002; Bach & Dmitrenko, 2003). Consequently, the electrophilicity at this position is increased and hydride transfer is enhanced, provided that substrates are also suitably oriented and parallel to the plane of the isoalloxazine ring. Third, several gas-phase theoretical studies and crystallographic studies of free flavins and flavin analogs have shown that the two electron reduced isoalloxazine ring is "bent" by up to 30° along the N5-N10 axis (Guo *et al.*, 2002; Rizzo, 2001; Reibenspies *et al.*, 2000). Such a "butterfly" conformational

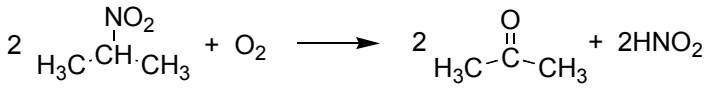
change associated with redox perturbation should significantly influence catalysis in flavoproteins. Moreover, reaction with O_2 at C4a, or other acceptor molecules at either N5 or C4a, can then occur from either the *r*e-face or the *s*i-face. However, there are structurally defined examples where oxidized flavoenzymes have significantly bent isoalloxazine rings (Satoh *et al.*, 2003; Binda *et al.*, 2003; Binda *et al.*, 2002a; Binda *et al.*, 2002b) and other examples where reduced flavoproteins contain almost planar cofactors (Gatti *et al.*, 1994; Schreuder *et al.*, 1992; Bach & Dmitrenko, 2004; Khan *et al.*, 2004; Ortiz-Maldonado *et al.*, 2003; Khan *et al.*, 2002; Barna *et al.*, 2001; Ortiz-Maldonado *et al.*, 1999). The protein environment is thought to play a significant role in stabilizing these "higher energy" states. Thus, comprehensive rules that yield a given reaction coordinate remain to be defined in most flavoproteins.

2.4 Identification of Nitroalkane Oxidase as a Flavoprotein

[**]The ability of organisms to oxidize nitroalkanes to the respective aldehyde or ketone plus nitrite has been known for several decades (Little, 1951). The activity is found in both bacteria and eukaryotes; the latter are predominantly yeast and fungi (Kido *et al.*, 1975), although there is a report of a plant 3-nitropropionate oxidase (Hipkin *et al.*, 1999). Most of these enzymes are more active on secondary nitroalkanes than on primary nitroalkanes (Kido *et al.*, 1975). Several such enzymes have been purified (Kido *et al.*, 1976b; Gorlatova *et al.*, 1998; Dhawale, M. R., and & Hornemann, U., 1979; Kido *et al.*, 1978a); in each case, nitropropane is the best substrate and the stoichiometry of the

[**] The sections on NAO in this chapter are reformatted from the review article published by (Fitzpatrick *et al.*, 2005)

reaction is as shown in Scheme 2.1, so that these enzymes are classified as 2-nitropropane dioxygenases. The enzymes from *Hansenula mrakii* and *Neurospora crassa* are both flavoproteins (Kido *et al.*, 1976b; Gorlatova *et al.*, 1998; Francis *et al.*, 2005) and presumably so too are the others.

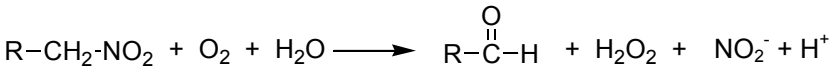


Scheme 2.1

Nitroalkanes are unusual acids, in that the anion protonates very slowly even at neutral pH and the nitro moiety lowers the pK_a value of the α-carbon to below 11 (Lewis, 1982), so that the anion can be used as a substrate at neutral pH. As a result, it has been possible to show that nitroalkane dioxygenase from *Neurospora crassa* is more specific for nitronates of 2-nitropropane and primary nitroalkanes with alkyl chain lengths between 2 and 6 carbons rather than the neutral substrates. With neutral nitroalkanes and anionic nitronates other than propyl-1- and propyl-2-nitronate, for which a non-enzymatic free radical reaction involving superoxide was established using superoxide dismutase, substrate oxidation occurs within the enzyme active site (Francis *et al.*, 2005).

In 1978, Kido *et al.,* described an enzyme from *Fusarium oxysporum* which oxidized a number of primary and secondary nitroalkanes (Kido *et al.*, 1978b). In contrast to the 2-nitropropane dioxygenases, this enzyme was more active on primary nitroalkanes. In addition, the reaction stoichiometry (Scheme 2.2) indicated that the

enzyme is properly termed a nitroalkane oxidase (NAO). NAO was present at much higher levels when the cells were grown in the presence of nitroethane, suggesting that nitroethane is a physiological substrate (Kido *et al.*, 1978b). While the enzyme was reported to have an absolute requirement for FAD for activity (Kurihara *et al.*, 1994), suggesting that it was a flavoprotein, the visible absorbance spectrum showed no evidence of the typical flavin absorbance (Figure 2.4). However, if the protein was denatured at neutral pH, a stoichiometric amount of FAD was released (Kurihara *et al.*, 1994; Gadda *et al.*, 1997).



Scheme 2.1

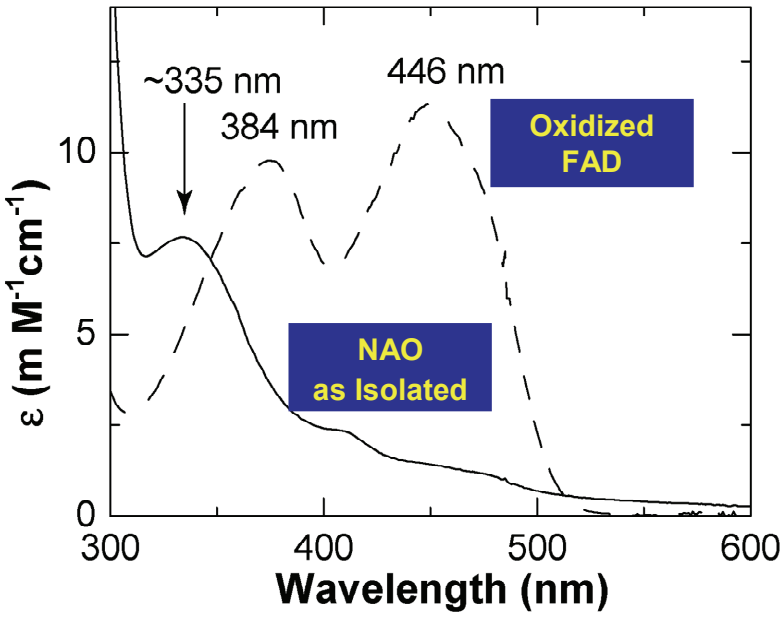


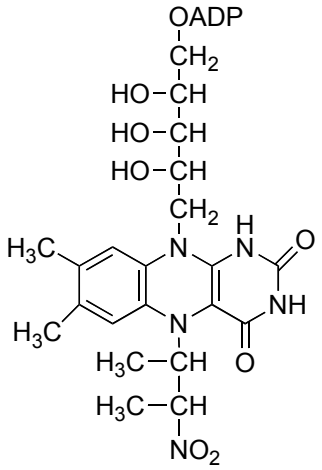
Figure 2.4 UV-Visible spectra of NAO as isolated from *Fusarium oxysporum* in contrast to spectra of oxidized FAD (Gadda *et al.*, 1997)

The structure of the cofactor in NAO as isolated from the fungus was determined by mass spectrometry, taking advantage of the stability of the protein-free cofactor at low pH. This showed two species with m/z^- values of 784.1415 ± 0.0068 and 887.2138 ± 0.0078. The smaller was identified as FAD, which has a molecular weight of 784; this was presumably formed during the laser desorption ionization process. The difference of 103.0645 ± 0.0103 between the two species established the cofactor as FAD plus $C_4H_9O_2N$; the added mass is equivalent to two nitroethane molecules minus nitrite. Consequently, the structure in Scheme 2.3 was proposed for the cofactor. When NAO containing this modified FAD was incubated at 37 ^oC, there was a slow conversion to an enzyme with a more typical flavoprotein spectrum. The activity of the enzyme showed a simultaneous large increase in activity, establishing that the cofactor adduct is not the active form of the cofactor (Gadda *et al.*, 1997). The FAD in the active enzyme could be removed reversibly and quantitatively to yield apoenzyme devoid of activity. This could be reconstituted with 5-deaza-FAD; the enzyme containing the modified flavin has no catalytic activity, and nitroethane does not reduce the flavin (Gadda & Fitzpatrick, 1998). These results established that active NAO contains unmodified FAD as the cofactor and suggested that the adduct is formed in the fungal cell during incubation with nitroethane.

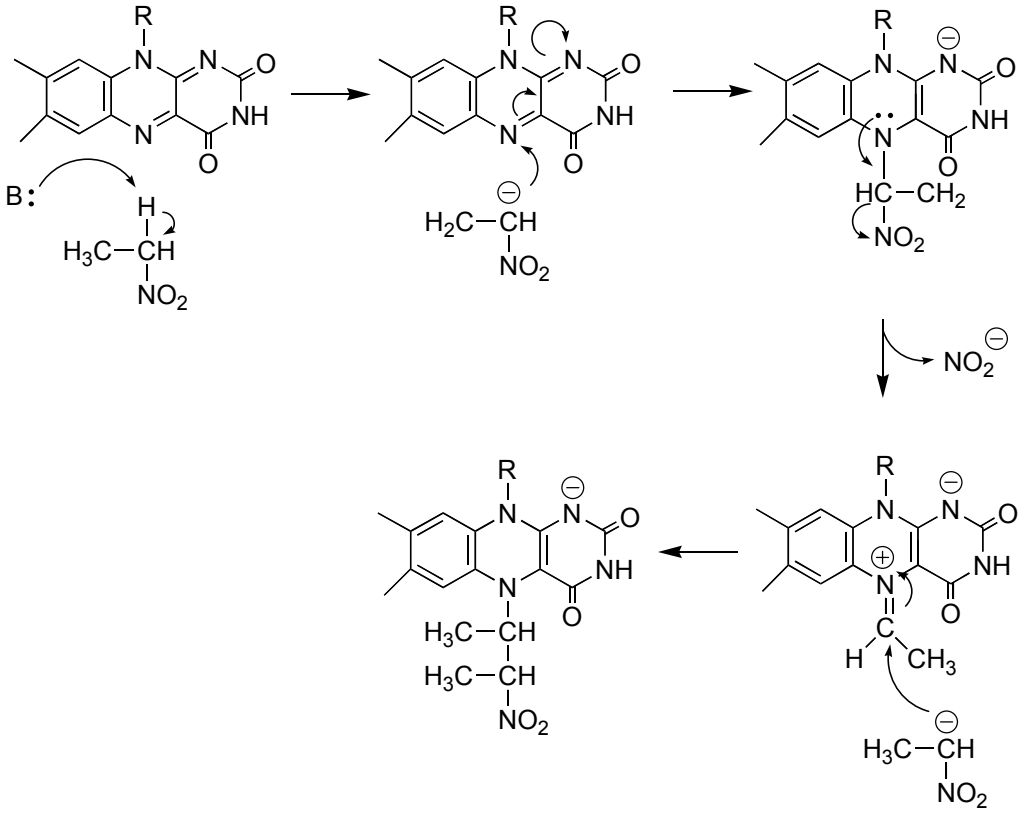
2.5 Mechanistic Studies on Nitroalkane Oxidase

Consistent with a redox role for the flavin cofactor in NAO, addition of neutral nitroethane to the FAD-containing form of the enzyme under anaerobic conditions results in an enzyme with a visible absorbance spectrum typical of reduced flavin; upon removal

of the substrate by aerobic gel filtration, the spectrum returns to that of oxidized flavin, with no evidence for any flavin N5-adduct. In contrast, if the enzyme is incubated with a mixture of neutral and anionic nitroethane, the resulting flavin is in the form of the adduct (Gadda *et al.*, 1997). This result is consistent with the mechanism of Scheme 2.4 for the formation of the adduct and with the neutral nitroalkane being the normal substrate. Formation of such an adduct would require one molecule of neutral nitroethane and one of the anion. Since the pK_a of nitroethane is 8.5, addition of nitroethane to cells at pH 8.0 would result in a mixture of the two forms in the growth media. Precedent for the mechanism of Scheme 2.4 comes from the reaction of D-amino acid oxidase with nitroethane anion; in that case a flavin adduct was trapped in the presence of cyanide and a similar mechanism was proposed (Porter *et al.*, 1973).



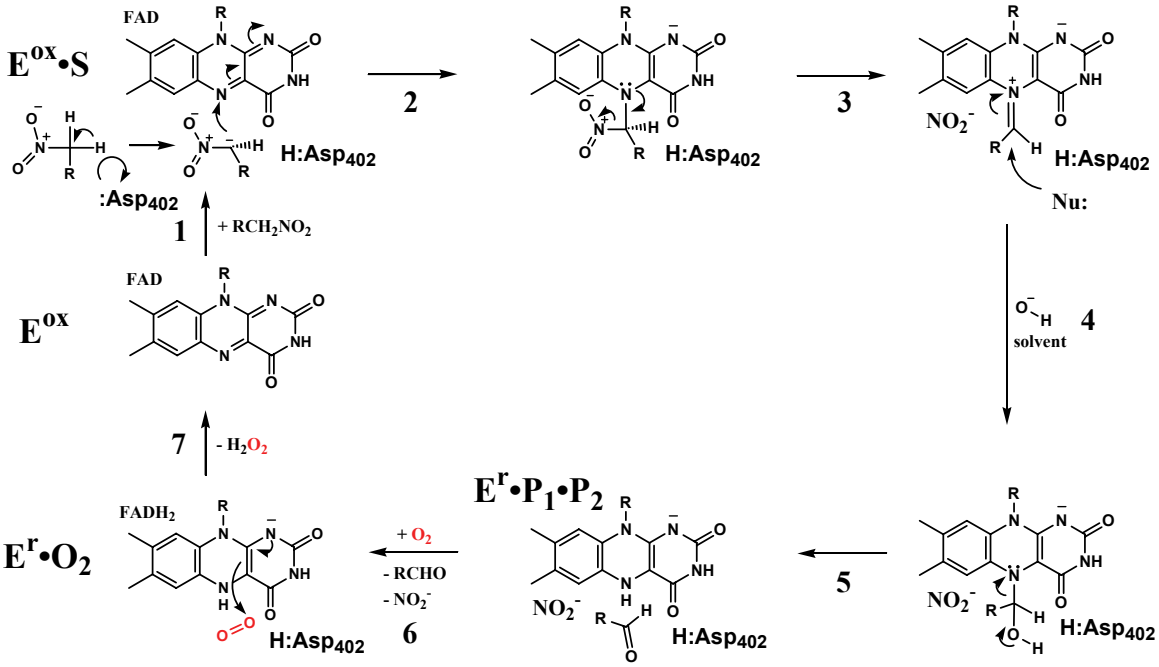
Scheme 2.3



Scheme 2.4

Thus, NAO differs from the 2-nitropropane dioxygenases in that catalysis involves the neutral nitroalkane. This preference of NAO for the neutral nitroalkane as substrate was confirmed by examining the k_{cat}/K_m pH profile for nitroalkane oxidase with a series of nitroalkanes with different pK_a values; if the substrate is allowed to reach an equilibrium protonation state at each pH before beginning the assay, the pK_a for basic limb of the bell-shaped pH profile matches that of the substrate (Heasley & Fitzpatrick, 1996). Since the initial steps in the formation of the flavin adduct involve the neutral nitroalkane, it is a reasonable assumption that these steps are along the normal catalytic pathway. Scheme 2.5 provides a mechanism consistent with these observations. Both

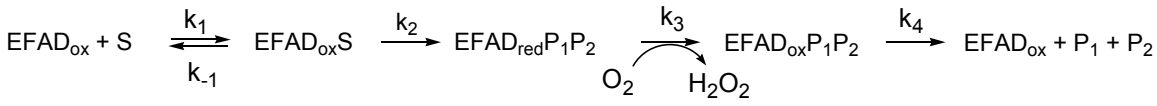
catalysis and inactivation begin in the same fashion, with proton abstraction from the neutral nitroalkane by an active site base. Nucleophilic attack of the nitroalkane anion on the flavin followed by nitrite loss would form an electrophilic cation. Attack by hydroxide rather than a nitroalkane anion on the cation would form a species which could eliminate the respective aldehyde to form reduced FAD. This provides a reasonable mechanism for oxidation of neutral nitroalkanes by NAO.



Scheme 2.5 (Nagpal *et al.*, 2005)

The kinetic mechanism of NAO with nitroethane and 1-nitrobutane as substrates has been examined using single turnover methods (Valley & Fitzpatrick, 2003a). The results are consistent with the mechanism of Scheme 2.6, where P_1 and P_2 are the products nitrite and acetaldehyde. This mechanism is similar to that seen with most

flavoprotein oxidases, in that flavin reduction occurs prior to the reaction with oxygen (Palmer & Massey, 1968; Bright & Porter, 1975) With nitroethane as substrate, cleavage of the substrate CH bond is rate-limiting for reduction, based on comparison of the isotope effects with [1,1-$^2H_2$]nitroethane on the k_{cat}/K_m value for nitroethane and on the limiting rate of reduction of the flavin measured in the stopped-flow spectrophotometer. The isotope effect on k_{cat} is much smaller, establishing that product release or a conformational change limits k_{cat}.



Scheme 2.6

The best substrates for NAO are linear primary nitroalkanes (Gadda & Fitzpatrick, 1999). Introduction of a hydroxyl group decreases the k_{cat}/K_m value about 200-fold. Comparison of the effect of chain length on the k_{cat}/K_m value showed a monotonic increase with each additional methylene group from nitroethane to 1-nitrobutane increasing the k_{cat}/K_m value by 20-fold and no further increases with 1-nitropentane or 1-nitrohexane. This is consistent with a contribution of 2.6 kcal/mol for each additional methylene, suggesting that the active site of NAO is a hydrophobic channel. A similar contribution of each methylene group has been found for stearoyl acyl carrier protein Δ^9-desaturase (Haas & Fox, 1999).

The k_{cat}/K_m pH profile with nitroethane as substrate shows that an enzyme base with a pK_a value of 7.0 must be unprotonated for catalysis (Gadda & Fitzpatrick, 2000e). This is presumably the active site base in Schemes 2.4 and 2.5, a conclusion subsequently confirmed by site-directed mutagenesis (*vide infra*). An identical pK_a is seen in the K_i pH profile for the competitive inhibitor valerate; this result and the isotope effect on the k_{cat}/K_m value for nitroethane establish 7.0 as the intrinsic pK_a value of the base. With primary nitroalkanes of increasing chain length, the pK_a seen in the k_{cat}/K_m pH profile shifts to lower values and the deuterium isotope effect on the k_{cat}/K_m value decreases. For the kinetic mechanism of Scheme 2.6, both effects can be attributed to an increase in the forward commitment to catalysis, k_2/k_{-1}. This has allowed the effects of chain length on the K_d value, k_{-1}/k_1, and on catalysis, k_2, to be separated. As the chain length increases from nitromethane to 1-nitrobutane, each methylene group results in a decrease of the K_d value of 15-fold, equal to 1.7 kcal/mol, and an increase in k_2 of 4-fold . With nitroethane as substrate, the value of k_2 has been directly measured as 250 s^{-1} by stopped-flow methods, while that for 1-nitrobutane is at least 4000 s^{-1} (Valley & Fitzpatrick, 2003b).

As noted above, with nitroethane as substrate the rate-limiting step for flavin reduction is formation of the nitroethane anion. Accordingly, the acetate-catalyzed formation of nitroethane anion provides a nonenzymatic model for the reaction of NAO. The temperature dependences of both the enzymatic and the acetate catalyzed reactions were determined as a function of temperature for nitroethane and [1,1-$^2H_2$]nitroethane (Valley & Fitzpatrick, 2004). In the case of the enzyme, the temperature dependence of the k_{cat}/K_m value for nitroethane was determined, since this provides the appropriate

second order rate constant for comparison with the second order rate constant for the nonenzymatic reaction. Direct comparison of the rate constants for the enzymatic and nonenzymatic reactions gives a value for the rate acceleration by NAO of 1.2 billion-fold. This is likely to be an underestimation of the catalytic power of the enzyme, since the k_{cat}/K_m value for 1-nitrobutane is 400-fold that for nitroethane and is limited by binding rather than anion formation (Gadda & Fitzpatrick, 1999; Gadda *et al.*, 2000d), while the second order rate constant for nonenzymatic ionization of 1-nitrobutane is comparable to that for nitroethane (Gadda *et al.*, 2000d).

Analyses of the temperature dependence of the enzymatic and non-enzymatic reactions allowed calculation of the thermodynamic parameters for formation of the transition state in each case. The rate increase seen in the enzyme-catalyzed reaction could be attributed to a decrease in $\Delta H^{\ddagger}$ of 18.6 kcal/mol (Valley & Fitzpatrick, 2004), in agreement with the observation of Wolfenden that enzyme reactions typically exhibit large decreases in $\Delta H^{\ddagger}$ compared to the corresponding nonenzymatic reactions (Wolfenden *et al.*, 1999). More uniquely, analysis of the temperature dependence of the deuterium isotope effects on the rate constants could be used to probe the contribution of quantum mechanical tunneling to the reaction and to determine if there was a greater contribution of tunneling to the enzyme catalyzed reaction. Deuterium or tritium isotope effects greater than the semiclassical value, 7 or 16, respectively, are often taken as an indication of a contribution from tunneling. However, deuterium isotope effects significantly less than 7 can occur in reactions involving quantum mechanical behavior (Kohen & Klinman, 1999). A more reliable criterion for a contribution of tunneling to a

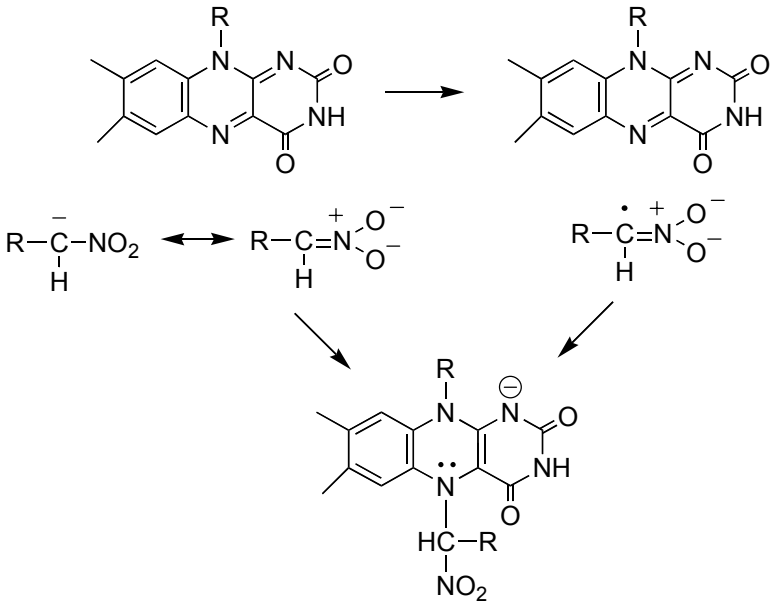
reaction is the isotope effect on the Arrhenius prefactor A. This is expected to have a value of unity in the absence of tunneling, although more realistic limits are 0.6-1.4 (Bell, 1974). Tunneling can result in isotope effects on A above or below this range, depending upon the magnitude of the contribution from tunneling to the reaction (Kohen & Klinman, 1999). In the case of NAO, the value of A_H/A_D is within the range expected in the absence of a significant tunneling contribution, 0.85 ± 0.35 (Valley & Fitzpatrick, 2004). More importantly, it is comparable to that for the nonenzymatic reaction, 0.89 ± 0.25. While this is consistent with a lack of contribution of tunneling to both reactions, it is more correct to conclude that tunneling contributes to an equivalent extent to the enzymatic and nonenzymatic reaction.

Because CH bond cleavage is rate-limiting with nitroethane, the best characterized substrate for NAO, the understanding of this step is better understood than the subsequent steps in catalysis. A critical question is whether the flavin adduct formed by attack of the nitroalkane anion during inactivation is also on the normal catalytic pathway, as proposed in Scheme 2.5. Cyanide has been used to trap this intermediate and establish its kinetic competency (Valley *et al.*, 2005). When assays are carried out with nitroethane in the presence of cyanide the enzyme rapidly loses activity. Rapid reaction analyses establish that cyanide can react with the enzyme in the presence of nitroethane at the same rate as nitroethane can reduce the flavin and that inactivation occurs before reaction of the reduced enzyme with oxygen. Use of 1,1-[$^2H_2$]nitroethane decreases the rate of inactivation by cyanide; the isotope effect equals the isotope effect on the k_{cat}/K_m value. These results are consistent with cyanide intercepting a species which is formed

during normal turnover after removal of the substrate proton but before formation of fully reduced flavin and aldehyde. Mass spectrometry of the flavin isolated from the cyanide-inactivated enzyme establishes that it is in the form of an N5-flavin adduct analogous to that of Scheme 2.3, but with a cyano moiety replacing the second nitroethyl group; i. e., 5-cyanoethyl-1,5-dihydro-FAD. This is consistent with cyanide inactivation involving a mechanism similar to that of Scheme 2.4, and establishes the intermediacy of the electrophilic flavin cation in Schemes 2.4 and 2.5 in normal catalysis.

A direct nucleophilic attack of the anion on the flavin is the most straightforward mechanism for formation of this adduct. However, in studies of the reactions of anions with flavins, the formation of N5 adducts is suggested to involve a radical intermediate (Scheme 2.7) (Bruice, 1980). While the mechanism of nitroethane anion addition to oxidized FAD has not been examined for NAO, it has been in the case of D-amino acid oxidase (DAAO). DAAO will not react with neutral nitroethane, but will catalyze the oxidation of nitroethane anion to acetaldehyde and nitrite (Porter *et al.*, 1973). Although a 5-nitrobutyl adduct is apparently not formed, if this reaction occurs in the presence of cyanide, a 5-cyanoethyl FAD adduct analogous to that in NAO and Scheme 2.3 can be isolated. While the DAAO reaction is certainly not physiological, it does serve as a model for the NAO reaction. When [1-^{2}H]nitroethane anion is used in the DAAO reaction, there is an isotope effect of 0.84 on the k_{cat}/K_m value (Kurtz & Fitzpatrick, 1997). This is consistent with a change in hybridization at the α-carbon from sp^2 to sp^3. In solution the nitroethane anion is the nitronate, so that the α-carbon is sp^2 hybridized (Bordwell & Boyle Jr., 1975). This carbon would similarly be sp^2 hybridized in the nitronate radical,

but sp^3 in the adduct. The inverse isotope effect in the DAAO reaction is most consistent with the radical not being an intermediate in formation of the adduct, but rather a mechanism involving direct nucleophilic attack of the nitronate on the flavin. By analogy, the NAO reaction would similarly involve a direct nucleophilic attack on the cofactor by the anion formed upon abstraction of the α-proton from the nitroalkane by the active site base, Asp402.



Scheme 2.7

2.6 Nitroalkane Oxidase, a New Member of Acyl-CoA Dehydrogenase Family

Using the sequences of tryptic peptides to generate oligonucleotide probes, the cDNA for nitroalkane oxidase was isolated from a *F. oxysporum* cDNA library of cells expressing the enzyme. The protein sequence derived from the cloned cDNA showed that the enzyme had homology to members of the acyl-CoA dehydrogenase (ACAD) family (Daubner *et al.*, 2002) . The identity extends throughout the entire sequence, and the regions of greatest divergence correspond to external loops in the structure of medium chain ACAD (MCAD). While this homology was unexpected, it does make chemical sense. Both enzymes are flavoproteins which catalyze the abstraction of an acidic proton as the first step in catalysis (Scheme 2.5, Figure 2.5), after which the mechanisms diverge. In the case of the MCAD, the base is Glu376 (Thorpe & Kim, 1995) but same role is played by Asp402 in NAO. Mutagenesis of Asp402 in NAO to glutamate decreases the rate of CH bond cleavage by the enzyme by 15 to 20-fold, while mutagenesis to asparagine or alanine decreases the rate at least 100-fold (Valley & Fitzpatrick, 2003a, 2003b). In addition, the k_{cat}/K_m pH profiles for the D402N and D402A enzyme no longer show a pK_a of 7 (Valley & Fitzpatrick, 2003a). This is consistent with Asp402 being the active site base with a pK_a of 7. The long lifetimes of nitroalkane anions in solution at neutral pH made it possible to more directly determine whether Asp402 was indeed the base responsible for abstracting a proton from the substrate to form the anion. At pH 8, the D402A enzyme has no significant activity above background with neutral nitroethane as substrate. However, use of nitroethane anion as the substrate results in recovery of substantial activity (Valley & Fitzpatrick, 2003a). This

rescue of the mutant with the anion provides strong support for the proposed role of Asp402 as the base in Scheme 2.5.

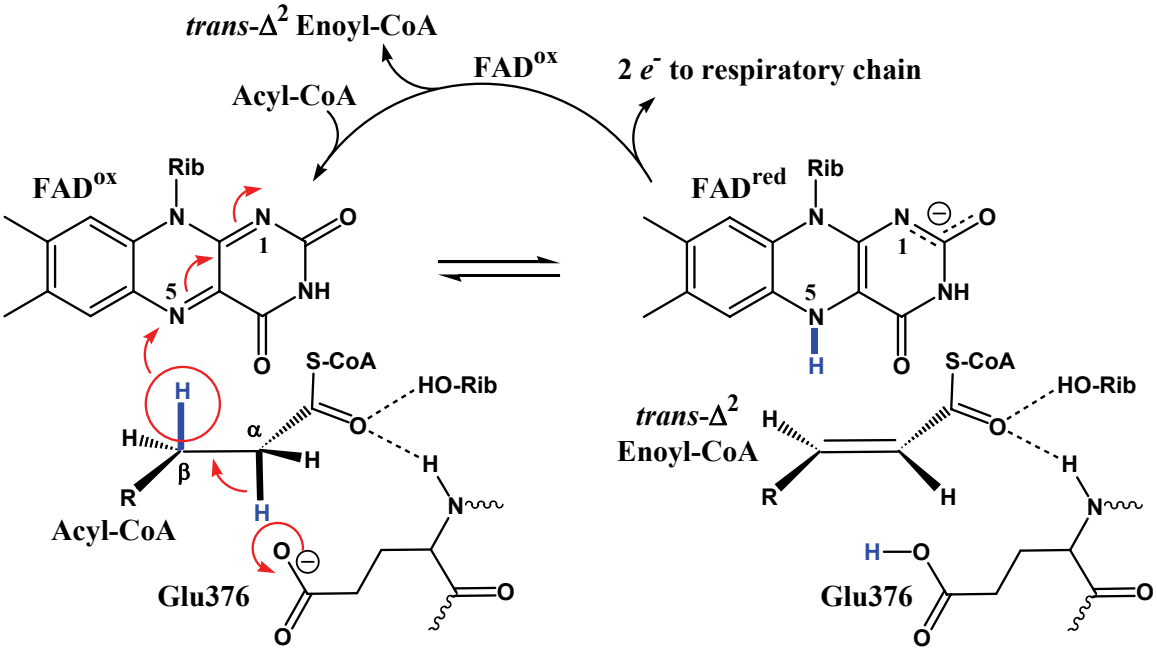


Figure 2.5 Proposed reaction mechanism for Acyl-CoA dehydrogenase (Ghisla & Thorpe, 2004).

Mechanistic studies of the ACAD enzymes were pioneered by Helmut Beinert almost five decades ago (Steyn-Parve & Beinert, 1958b, 1958a; Beinert & Page, 1957; Crane *et al.*, 1956a; Hauge *et al.*, 1956; Crane & Beinert, 1956b) and the research area is still very active. Indeed, the structural, mechanistic, mutagenic and spectroscopic analysis of the MCAD provides exquisitely detailed proposals for the structural contributions to catalysis (Kim & Miura, 2004; Ghisla & Thorpe, 2004). However, to our knowledge

none of the ACAD or ACO enzymes will transform nitroalkanes, nor will NAO transform acyl-CoA substrates (Daubner *et al.*, 2002). Members of the ACAD superfamily are roughly classified into four branches based upon substrate preference and reaction type (Ghisla & Thorpe, 2004; Kim & Miura, 2004; Tanaka & Indo, 1992). The first branch catalyzes the catabolism of fatty acids via β-oxidation. The chain length of the acyl-thioester substrate dictates which ACAD homolog catalyzes the α, β-dehydrogenation reaction to yield the corresponding *trans*-2,3-enoyl CoA product. Thus the short, medium, long and very long chain acyl-CoA dehydrogenases demonstrate preferences for C_4-C_8 (SCAD), C_4-C_{16} (MCAD), C_{10}-C_{16} (LCAD) and $>C_{16}$ (VLCAD), respectively. Because of the many prolific research groups focused on the MCADs, they are the best characterized of the whole superfamily (Ghisla & Thorpe, 2004). The second group of homologs is specific for branched acyl-thioester substrates encountered in amino acid catabolism. In mammals, the enzymes from the first two branches are located inside the mitochondrial matrix and transfer two reducing equivalents to the electron-transferring flavoprotein (ETF) (Crane *et al.*, 1956a). The reducing equivalents are subsequently passed on to the ETF-dehydrogenase and then to the respiratory chain for ATP production. Thus, these members of the ACAD superfamily evolved a strategy to prevent oxidation of $FADH^-$ by O_2 as discussed in two recent reviews (Ghisla & Thorpe, 2004; Kim & Miura, 2004). The third branch is a more distantly related group of peroxisomal enzymes which degrade longer fatty acid chains ($>C_{23}$). These enzymes transfer the reducing equivalents to oxygen to produce H_2O_2 and are thus acyl-CoA oxidases (ACO) (Setoyama *et al.*, 1995; Kunau *et al.*, 1995; Osumi *et al.*, 1980; Inestrosa *et al.*, 1979), (Pedersen & Henriksen, 2005). The fourth group of enzymes has only

recently emerged and is best characterized for the acyl-carried protein dehydrogenase (ACPD) from *Streptomyces hygroscopicus*. They function in the biosyntheses of the immunosuppressive and antifungal macrolides FK520 and FK506 (Wu *et al.*, 2000; Motamedi & Shafiee, 1998). These ACAD homologs are essential for the synthesis of the unique "glycolate" extender units that are incorporated into many polyketide backbones by modular polyketide synthetases (Carroll *et al.*, 2002; Cane *et al.*, 1998; Walsh, 2004). It seems unlikely that an ETF analog is essential in this enzyme system.

At least one crystal structure is available from each branch (Nagpal *et al.*, 2004; Kim & Miura, 2004; Battaile *et al.*, 2004; Watanabe *et al.*, 2003; Satoh *et al.*, 2003; Nakajima *et al.*, 2002; Lee *et al.*, 1996; Thorpe & Kim, 1995; Djordjevic *et al.*, 1995; Kim, J. J. *et al.*, 1993; Kim *et al.*, 1992; Kim & Wu, 1988; Battaile *et al.*, 2002; Pedersen & Henriksen, 2005). With amino acid sequence similarity to NAO as low as 7-9% for ACO and ACPD, respectively; potential homologs cannot be unambiguously identified from sequence alignment alone. Indeed, *BLAST* searches with NAO do not return ACO as a homolog.

The oxidative half-reactions of NAO and ACO differ from that of MCAD (Ghisla & Thorpe, 2004). In particular, reduced MCAD with *trans*-2,3-enoyl CoA bound does not react with O_2. Because the enoyl-CoA product binds to the reduced MCAD with more affinity, it does not dissociate readily. The tight product binding in case of ACADs prevents the O_2 access to the $FADH^-$ species, principally by steric exclusion. However, the reduced MCAD-(enoyl-CoA)-$FADH^-$ complex binds to the ETF protein. The

complex is hypothesized to transfer reducing equivalents in two one-electron transfer events, since the inter flavin distance in the complex does not permit direct contact between redox partners (Parker, 2003; Roberts *et al.*, 1996; Roberts *et al.*, 1999). A very recent crystal structure of the complex between MCAD-ETF supports this theory and furthermore suggests significant conformational flexibility is an essential feature of the complex (Toogood *et al.*, 2004). The affinity of oxidized MCAD for the enoyl-CoA product is greatly reduced, allowing it to dissociate. In contrast, NAO and ACO transfer their reducing equivalents to O_2 and produce H_2O_2. The sequence alignments suggest that Ser171 in NAO is analogous to Gly178 in ACO-II and to Gly177 in ACO-I, which may increase the accessibility of O_2 to $FADH^-$ species. Thus ACO may allow enoyl-CoA products to dissociate readily and/or allow oxygen access even in the product complex. However, the active site of NAO is much more similar to MCAD than ACO. The detailed crystal structure analysis of NAO and its reaction intermediates will help in understanding the structural basis for the different activities, as well as potential ancestral links to the MCAD, ACO and other superfamily members and redox partners.

3 MATERIALS & METHODS

This chapter details the experimental techniques used to solve the crystal structures of a substrate trapped N5-FAD adduct of NAO (ES*) and the oxidized form of NAO to 2.2 Å and 2.07 Å resolutions, respectively. X-ray structure determination of NAO in its different reaction states was not only of interest from a catalytic viewpoint, but also from a crystallographic viewpoint, owing to the challenges encountered during this project. To list a few of the challenges, numerous soaking and co-crystallization trials with a wide variety of heavy atoms to phase ES* were unsuccessful. The attempts of solving the structure with molecular replacement were also futile. After much struggle, the structure of ES* was finally solved with the SeMet MAD phases derived from the location and the refinement of 52 selenium sites. The large number of selenium sites in ES* gave us a chance to learn a great deal about different phasing programs, their capacities, and their limitations. To our knowledge, this is one of the few SeMet structures with such a large number of selenium sites.

Although the crystal structure of EI was solved by using ES* as a molecular replacement search model, the collection of native data in this case, pushed the limits of synchrotron facilities at Brookhaven and Argonne National Labs. The EI crystals grew with an exceptionally long *c* axis (485 Å) and a huge unit cell volume (4,570,000 $Å^3$). The crystal form was interesting and worth pursuing since it diffracted to a very high resolution (1.6 Å) and had a very low mosaic spread (0.3°). After 2 years of struggle and innumerable synchrotron trips, we were able to collect and process 2.07 Å resolution native data of EI. Out of total 31,644 protein structures deposited in PDB, only 17 crystal

structures have a *c* axis greater than 485 Å. None of the protein crystal structures with a *c* axis greater than 485 Å diffract to a resolution greater than 2.2 Å, which makes the crystal structure of EI, the highest resolution structure deposited in PDB.

3.1 X-ray Structure Determination of N5-(2-nitrobutyl)-1,5-dihydro-FAD Adduct of NAO (ES*)

3.1.1 Protein Expression and Purification of ES* and SeMet-ES*

Recombinant nitroalkane oxidase was expressed and purified from *E.coli* strain BL21 (DE3) transformed with pETNAO4 (Daubner *et al.*, 2002; Valley & Fitzpatrick, 2003b).The purified enzyme was stored in 20 mM HEPES pH 8.0,1 mM EDTA, 5% glycerol and aliquots were flash frozen with liquid N_2. SeMet-ES* was obtained by expression in the methionine auxotroph *E.coli* B834 (DE3) (Novagen) as previously described (Nagpal *et al.*, 2004). To generate the covalent adduct, 25 μM NAO was incubated with 28 mM neutral nitroethane and 9 mM anionic nitroethane in 500 mM HEPES, pH 8.0, for 1 hour at 25 °C. The enzyme was then exchanged into 20 mM HEPES, pH 8.0, 1 mM EDTA, 5% glycerol at 4 °C using an Ultrafree-15 centrifugal filter device (Millipore Co., Bedford, MA) and aliquots were flash frozen with liquid N_2. NAO has a subunit molecular weight 47,955 Da, each of which contains 13 methionine residues, excluding the N-terminal residue (Gadda & Fitzpatrick, 1998; Daubner *et al.*, 2002).

3.1.2 Crystallization of ES* and SeMet-ES*

Microcrystalline showers of the ES* were initially obtained from the flexible sparse matrix screen prepared according to (Zeelen *et al.*, 1994). Diffraction quality crystals were obtained at 4 °C by mixing 2 µl of 10 mg/ml protein solution (in 10 mM PIPES, pH 7.5) with 2 µl of reservoir solution containing 20% w/v PEG 1500, 300 mM ammonium sulphate in 150 mM PIPES (pH 7.5) (Figure 3.1 *a*). The crystals typically grew within 20 days with the hanging drop, vapor diffusion method. Single crystals of SeMet-ES* were grown from similar conditions except for the addition of 10 mM tris(2-carboxyethyl) phosphine hydrochloride (TCEP-HCl) to the crystallization drop. Crystals in natural and SeMet enriched forms of ES* grew in orthorhombic space group with similar cell edges (a = 90.5 Å; b = 163.9 Å; c = 172.3 Å).

3.1.3 X-ray Data Collection and Processing of Native ES*

Single crystals were mounted in nylon loops (no cryoprotectant was necessary) and flash frozen by quick submersion into liquid N_2. All X-ray diffraction data were collected from crystals held at approximately 100 K. The native dataset for ES* was collected with 180° phi rotation at beamline X26C of the NSLS using 1.0 Å x-rays and an ADSC Quantum 4 detector. Each image was collected with 60 sec exposure, 0.5° oscillation and 200 mm crystal to detector distance (Figure 3.1 *b, c*). The crystals diffracted to 2.3 Å resolution with 99.9 % completion in the highest resolution shell.

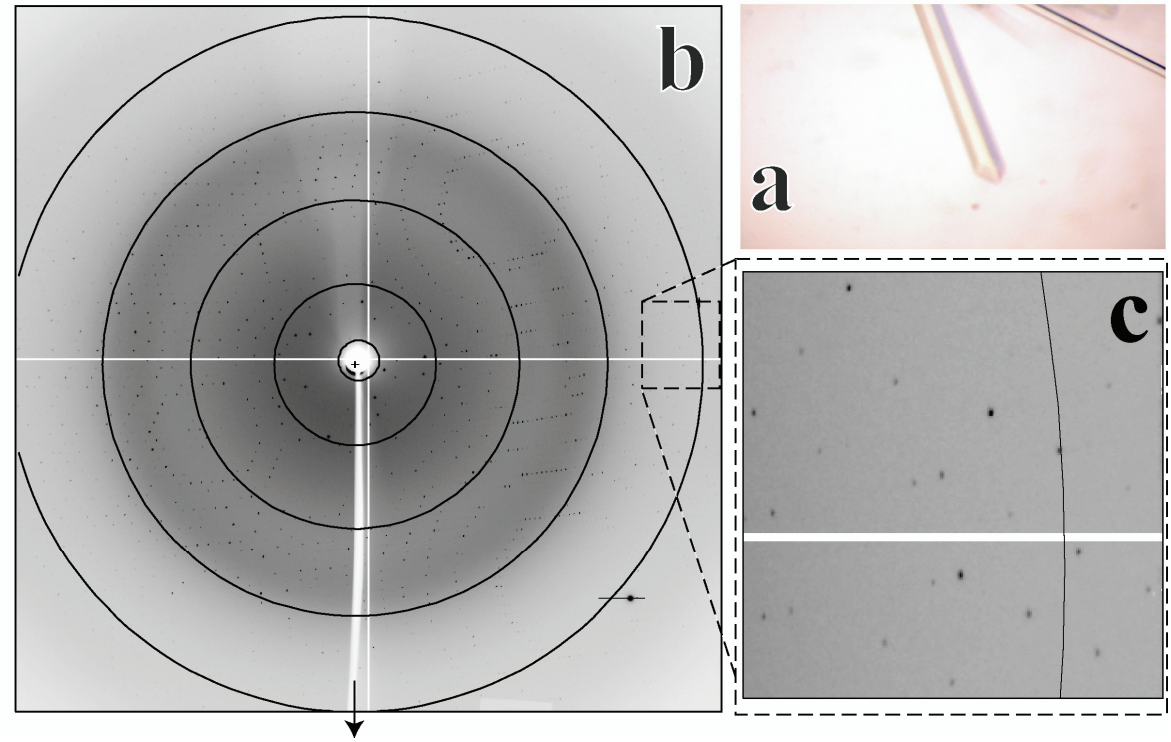


Figure 3.1 a) Crystals of the N5-FAD adduct of nitroalkane oxidase photographed without polarization (0.3 X 0.09 mm), b) The x-ray diffraction pattern at 100K with rotation about the vertical axis collected with ADSC Quantum 4 detector at the NSLS (beamline X26C) with 60 s exposure, 0.5° oscillation and 200 mm crystal to detector distance, arcs indicate 2.4, 3.2, 4.7, 7.8 Å. c) An expanded and contrast adjusted view of the high resolution diffraction perpendicular to the rotation axis from (b).

3.1.4 Phase Problem in Crystallography (Background)

In the standard crystallographic experiment a crystal is positioned in a beam of monochromatic X-ray radiation. X-rays passing through the crystal will cause the electrons of the molecules to oscillate. These oscillating charges then emit X-ray radiation of the same wavelength in all directions. A crystal consists of a periodic arrangement of molecules where repeating units, the so-called unit cells, form a three-dimensional lattice. Because of this periodicity, the waves scattered by the atoms in all

unit cells will only interfere constructively in certain discrete directions. This leads to the Laue diffraction conditions (3.1):

$$\begin{aligned} \vec{a} \,.\, \vec{S} &= h_1 \\ \vec{b} \,.\, \vec{S} &= h_2 \\ \vec{c} \,.\, \vec{S} &= h_3 \end{aligned} \tag{3.1}$$

where $\vec{a}$, $\vec{b}$, and $\vec{c}$ are the translation vectors of the crystal lattice, Laue-indices h_1, h_2, and h_3 are integers and $\vec{S}$ is called the diffraction vector. An alternative way to describe diffraction is to consider a diffracted beam as being reflected by a plane *hkl* through the endpoints of vectors $\vec{a}$ / h, $\vec{b}$ / *k* and $\vec{c}$ / l. (Miller indices h, k, and *l* are related to the Laue indices: $h_1 = n$h, $h_2 = nk$ and $h_3 = n$l, *n* being an integer). Diffraction only occurs if the angle θ of the incident beam with the lattice plane *hkl* satisfies Bragg's law (3.2):

$$2\, d_{hkl} \sin(\theta) = n\, \lambda \tag{3.2}$$

where d_{hkl} is the distance between adjacent lattice planes *hkl* and λ is the wavelength of the incident beam; integer *n* is called the order of the reflection. The direction of diffraction vector $\vec{S}\,(hkl)$ is normal to the reflecting plane *hkl* and its length $|\vec{S}\,(hkl)|$ is equal to $1/d_{hkl}$. The endpoints of all vectors $\vec{S}\,(hkl)$ form a three-dimensional lattice with translation vectors $\vec{a}^*$, $\vec{b}^*$ and $\vec{c}^*$ (with $\vec{a}^* = \vec{b} \times \vec{c}/V$, $\vec{b}^* = \vec{c} \times \vec{a}/V$, and

$\vec{c}^{\,*} = \vec{a} \times \vec{b} / V$, where V is the volume of the unit cell), the so-called reciprocal lattice. This lattice allows $\vec{S}(hkl)$ to be calculated in a convenient way from (3.3):

$$\vec{S}(hkl) = h\,\vec{a}^{\,*} + k\,\vec{b}^{\,*} + l\,\vec{c}^{\,*} \tag{3.3}$$

The experimentally measured intensity of reflection *hkl* depends on the distribution of the electrons in the unit cell. The electron density, ρ (*xyz*) is proportional to the square of the amplitude of structure factor F (*hk*l), see equation 3.4:

$$F(hkl) = V \int_x \int_y \int_z \rho(xyz)\, \exp\{2\pi i\,(hx + ky + lz)\}\, dx\, dy\, dz \tag{3.4}$$

with x, *y* and *z* being fractional coordinates. By inverse Fourier transform, the electron density distribution in the unit cell is calculated from structure factors F (*hk*l) as follows:

$$\rho(xyz) = 1/V \sum_h \sum_k \sum_l F(hkl) \exp\{-2\pi i\,(hx + ky + lz)\} \tag{3.5}$$

From the electron density distribution, an atomic model of the molecules in the unit cell can be constructed. However, the structure factor, F (*hk*l), in equation 3.6 is complex quantity with an amplitude and a phase:

$$F(hkl) = \left|F(hkl)\right| \exp\left(i\varphi(hkl)\right) \tag{3.6}$$

From the standard monochromatic experiment, amplitude F (*hk*l) can be derived from the measured intensity, but all information about phases $\varphi(hkl)$ is lost. Therefore, the electron density distribution cannot be constructed directly using equation 3.5. This problem is known as the crystallographic phase problem.

3.1.4.1 Ab initio Phasing Methods (Background)

In *ab initio* phasing, phase estimates are obtained from a single set of structure factor amplitudes, without using any of the experimentally determined intensity differences. In this case, the phase problem may be overcome by incorporation of additional, *a priori* available knowledge.

1) Direct Methods

Even the simplest form of prior knowledge, the expectation that the electron density is non-negative and consists of separated atoms throughout the unit cell, leads to statistical relationships among the structure factors that can be used to solve the phase problem. In the 1950's, Karle and Hauptman defined a functional form to express these relationships and thereby opened the field of direct methods (Karle & Hauptman, 1950). With recent advancements, direct methods are routinely used to solve thousands of structures consisting of up to 250 non-hydrogen atoms every year. The use of direct methods in protein crystallography has been limited because of the requirement of atomic resolution X-ray data and low phase reliability for the large number of atoms in the unit cell (Hauptman, 1997). Recently, programs like Shake-and-Bake have been used to solve substructures containing 2000 non-hydrogen atoms but high resolution data is still a

requirement. The program differs from the conventional direct method procedures in repetitively and unconditionally alternating the reciprocal-space phase refinement (shaking) with density modification (baking), which is then used to impose phase constraints in real space (Weeks *et al.*, 1993; Sheldrick & Gould, 1995) .

2) Low Resolution Phasing

Several attempts have been made to solve the phase problem by first locating the protein molecular envelope in the unit cell by only using the lowest resolution reflections. Many selection criteria based on electron density histograms and connectivity properties, statistical likelihood, atomicity etc. have been developed to select the best set of phases from a large number of random trial sets (Podjarny & Urzhumtsev, 1997). A common feature of such criteria is that they do not unambiguously judge the phase quality at low resolution. Nevertheless, the selection of the phase sets with best criterion values increases the ratio of good phase sets in the ensemble considered. An approximate solution of the phase problem may then be found by averaging the selected phase sets. Cluster analysis of the selected phase sets and averaging within clusters allow further improvement of this solution (Lunin *et al.*, 1998). These methods are important where the knowledge of the low resolution envelope is needed to solve the high resolution phase problem. For example, in molecular replacement, the approximate knowledge of the position and orientation of the molecule is helpful (Lunin *et al.*, 2000). The low resolution *ab initio* methods have also proved helpful in obtaining the low resolution image of the large molecular complexes. For example, these methods were used to phase the 50S ribosomal subunit from *Thermus thermophilus* to 40 Å resolution (Urzhumtsev *et*

al., 1996). The low resolution methods may also be important for crystals, which do not diffract to high resolutions especially those of viruses (Lunin *et al.*, 2000). Recently, low resolution methods have been coupled with averaging and density modification methods to further improve the quality of phases (Urzhumtsev *et al.*, 2000).

3) Molecular Replacement Method

A much more common method of *'ab initio'* phasing in protein crystallography is molecular replacement. In molecular replacement (Rossmann, 2001) a known structure of the homologous protein having high sequence identity (30% or greater) is used as prior phase information. Mathematically, the structure factor $\vec{F}_{hkl}$ is characterized by an amplitude $|\vec{F}_{hkl}|$ and a phase angle φ_{hkl}. For an unknown structure, the amplitude can be obtained by measurement of the reflection intensities from the X-ray diffraction data (3.7):

$$I_{hkl,\,obs} = |\vec{F}_{hkl,obs}|^2 = F_{obs}{}^2 \qquad 3.7$$

The calculated structure factor from the search model is given by:

$$\vec{F}_{calc} = \vec{F}_{hkl,calc} = \sum_j f_j \, e^{2\pi i\,(hx_j + ky_j + lz_j)} \qquad 3.8$$

where f_j is an atomic scattering factor.

As always the structure factor is characterized by an amplitude and a phase see equation 3.9;

$$\vec{F}_{calc} = \left| \vec{F}_{calc} \right| e^{i \varphi_{calc}} \qquad 3.9$$

Combining observed structure factors $|\vec{F}_{hkl,obs}|$ form equation 3.7 and phase angles φ_{calc} from equations 3.8 and 3.9, an electron density map for the biomolecule of interest can be calculated as follows:

$$\rho(x, y, z) = \frac{1}{V} \sum\sum\sum F_{obs} \; e^{-i[2\pi(hx + ky + lz) - \varphi_{calc}]} \qquad 3.10$$

Molecular replacement phasing algorithms obtain phase information by correctly placing the known model in the crystal lattice of the unknown structure. Traditionally, this problem was broken down into two three-dimensional search problems. Using Patterson search methods, first the correct orientation is determined, followed by a search for the correct translation vector (Rossmann & Blow, 1962). Recently, programs performing complete (Sheriff *et al.*, 1999) or directed (Kissinger *et al.*, 1999; Glykos & Kokkinidis, 2001) six-dimensional searches have been developed. Another recent development is the application of maximum likelihood to molecular replacement, which may allow the positioning of molecules with significantly lower homology (Randy, 2001). In the case of weak phasing power of the molecular replacement solution, model bias is a major problem. Methods like SIGMAA (Read, 1986) and simulated annealing in

combination with maximum likelihood refinement (Adams *et al.*, 1999) are helpful in removing model bias to some extent. More recent developments like prime and switch phasing in RESOLVE are proven to remove model bias. The method uses statistical density modification to remove model bias without including the phase information from the model, which was used for initial map calculations (Terwilliger, 2004).

3.1.4.2 Experimental Phasing (Background)

1) Single / Multiple Isomorphous Replacement Methods (SIR/MIR)

In macromolecular crystallography, it is common practice to solve the phase problem using additional experimental information (Drenth, 1999). In SIR or MIR methods, protein crystals are soaked in one or more solutions containing 'heavy' atoms. In the optimal case, this leads to specific binding of the heavy atoms to the protein molecules but the soaked crystals remain isomorphous to the native crystal. The differences in scattered intensities of the native and soaked crystals are significantly contributed from scattering of the heavy atoms and these differences can be used (for instance) to compute a Patterson map. Because there are only a few heavy atoms, such a Patterson map is relatively simple and easy to deconvolute. Alternatively, direct methods can also be applied to the intensity differences. Once the location of the heavy atoms in the crystal is determined, their contribution (F_{H1}) to the structure factors of the isomorphous derivative crystal (F_{PH1}) can be computed. If heavy atom binding does not alter the protein structure then F_{PH1} can be described as follows:

$$F_{PH1} = F_P + F_{H1} \quad 3.11$$

In the above equation phase and magnitude of F_{H1} are known but for F_{PH} and F_P only the magnitudes are known. In the Harker construction (Figure 3.2), the circle in green represents all the possible phases for F_P (structure factor for native protein). The radius of the circle is equal to $|F_P|$. After the first derivative is made with heavy atom H1, the possible phases for the derivatized protein crystals F_{PH1} can be represented by red circle. The structure factor from native protein F_P, has two possible solutions, a vector pointing towards A or vector pointing towards B (where the circle in red with the radius of $| F_{PH1}|$ intersects with the green circle). The phases for the second derivative can be represented by the circle in cyan, which has radius equal to $| F_{PH2}|$. Combining the phase information from both the derivatives, the phase angle for the unknown protein can then be unambiguously calculated. The final vector for protein of interest is represented by RF_P, where both the circles (in red and cyan) intersect at a common point A (Blundell & Johnson, 1976) (Figure 3.2). Since phases are derived by analyzing more than one derivative the method is called as multiple isomorphous replacement.

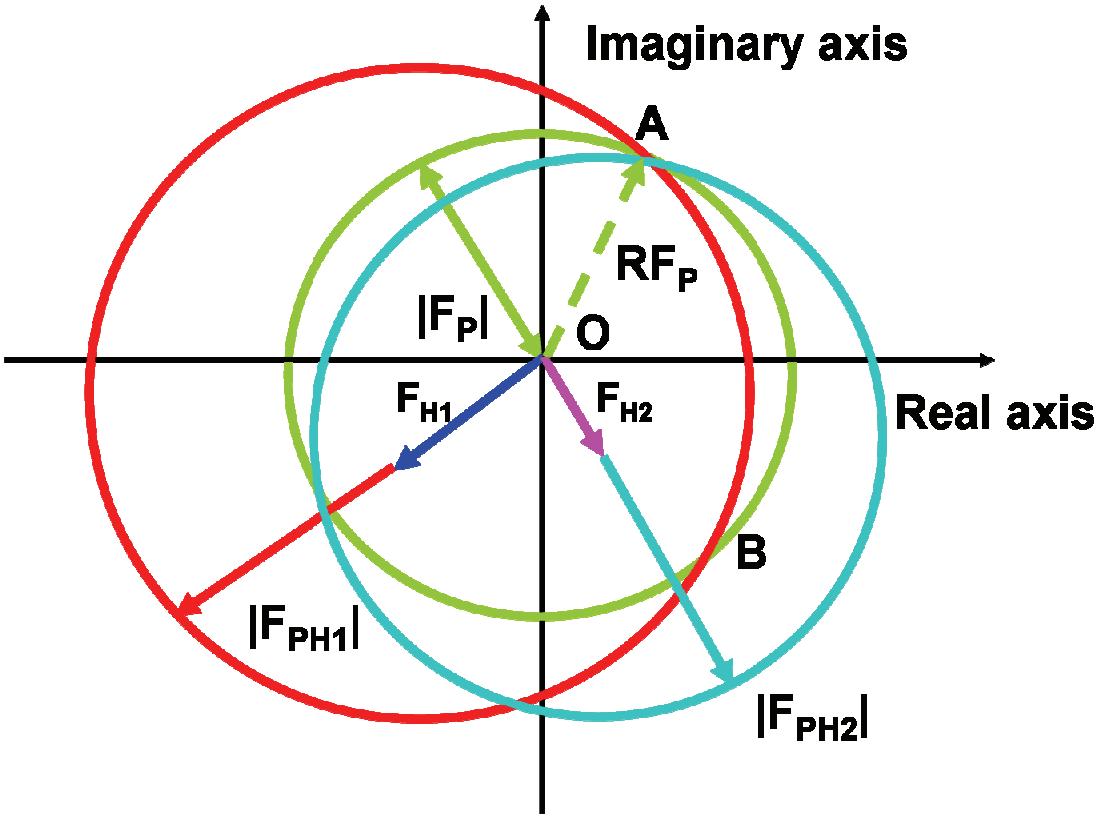


Figure 3.2 Harker construction for MIR method

2) Multiple /Single Wavelength Anomalous Dispersion

Many heavy atoms absorb X-ray radiation at wavelengths that are typically used in protein crystallography (0.6 - 2.0 Å). The absorbance of X-ray photons gives rise to anomalous diffraction. In anomalous scattering methods, the intensity differences between Friedel pair reflections *hkl* and *-h -k -l* (the so-called Bijvoet differences) at a single wavelength (SAD) can be used to calculate phase estimates.

The continuous tunability of synchrotron radiation sources has made it convenient to exploit yet another signal: dispersive intensity differences between data collected at

different wavelengths. In multiple-wavelength anomalous dispersion methods, MAD (Hendrickson & Ogata, 1997), a combination of the anomalous and dispersive signals allows phase determination from a single crystal. The basic idea is to locate the anomalous scattering atoms within the unit cell (those contributing to F_A) and use their position to calculate the corresponding phase angle φ_A. The MAD phasing equations (Karle, 1980) can then be used to generate an estimate for $\Delta\varphi$ (difference in phase angle between normal and anomalous scattering components) and $\mathbf{F_T}$ (the normal scattering component of all the atoms). In the simplest case this estimate is then used to the calculate phase of the $\mathbf{F_T}$, which is given as $\Delta\varphi + \varphi_A$. A Fourier transform of the amplitudes F_T and phases ($\Delta\varphi + \varphi_A$) yields an electron density map corresponding to all atoms in the structure (Figure 3.3).

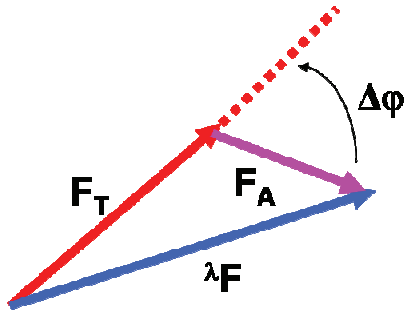


Figure 3.3 [††] **The vector diagram for calculating phases from MAD data set.** $^{\lambda}\mathbf{F}$ **represents the total observed scattering amplitude from the diffraction measurement at wavelength** λ**, which is unknown.** $\mathbf{F_T}$ **is the normal scattering component of all atoms whereas** $\mathbf{F_A}$ **is the anomalous scattering contribution from all atoms.** $\Delta\varphi$ **denotes the difference in phase angle between normal and anomalous scattering components.**

[††] Figure adapted from http://www.bmsc.washington.edu/scatter/MAD_1.html

MAD phasing has become increasingly popular due to the possibility to biosynthetically introduce anomalous scatterers into the protein itself (Ogata, 1998). In particular the *Eschericia coli* bacterium is used to substitute methionine for seleniomethionine. This yields anomalously scattering molecules with essentially equal structural properties compared to the native protein. Recently, density modification methods have been applied successfully to substitute the need for dispersive signals, allowing phasing by anomalous scattering using only a single wavelength, SAD (Wang, 1985).

3.1.5 Molecular Replacement Efforts to Solve ES*

cDNA analysis of NAO identified it as a new member of acyl-CoA dehydrogenase (ACAD) superfamily (Daubner *et al.*, 2002). The ACAD superfamily consists of different chain length acyl-CoA dehydrogenases (Kim & Miura, 2004), acyl-CoA oxidase I and II (ACOs) (Pedersen & Henriksen, 2005; Nakajima *et al.*, 2002) and acyl-ACP dehydrogenase (Watanabe *et al.*, 2003). NAO shares overall low sequence identity (21- 25%) with ACADs and BLAST searches did not return ACOs as possible homologs. ClustalW alignments showed that NAO only has 7 - 9% similarity with ACOs (Fitzpatrick *et al.*, 2005). Since molecular replacement requires high sequence identity protein homologs, medium chain acyl-CoA dehydrogenases (MCAD), which share 25% sequence identity with NAO were chosen as molecular replacement search models. Cross-rotation searches were performed with polyalanine models of the monomer as well as a tetramer of pig and human MCAD using CNS v1.1 (Brünger *et al.*, 1998a). Low resolution data between 15 to 4 Å resolutions was used to search for the NAO molecular

envelop in different orthorhombic space groups (*P*222, $P222_1$, $P2_12_12$ $P2_12_12_1$). Translation searches were performed using the top ten hits from the cross-rotation results. The searches gave a maximum correlation co-efficient of 15% and a minimum correlation co-efficient of 4.8% depending upon the molecular search model used. Since none of the molecular replacement results yielded reasonable correlation co-efficient or R_{factor}, phasing efforts were shifted to experimental phasing techniques for solving the three dimensional structure of ES*.

3.1.6 MIR Efforts to Solve ES* Crystal Structure

In efforts to use MIR methods for phasing ES*, the native protein crystals of ES* were derivatized with Au and Pt salts. The crystals of the ES* derivatives diffracted to a maximum resolution of 2.5 Å and were also isomorphous, as suggested by the R_{iso} values. The position of heavy atoms calculated using SOLVE v. 2.0.3 (Terwilliger & Eisenberg, 1983; Terwilliger *et al.*, 1987) did not yield interpretable maps, which led to the conclusion that heavy atom sites were not localized. Since MIR trials for ES* were not successful, the data collection and analysis details are not included in this chapter.

3.1.7 MAD Data Collection and Processing of SeMet-ES* Crystals

SeMet-ES* crystals were mounted in a nylon loop and flash frozen in liquid nitrogen. Fluorescence scans of a high diffraction quality crystal were collected at the **K** absorption edge of selenium. An optimized scan was processed using *CHOOCH* (Evans & Pettifer, 2001) and experimental f' and f'' (scattering factor and its imaginary part) plots were

calculated as a function of X-ray energy (Figure 3.4). A three wavelength SeMet MAD dataset collection strategy was designed based on f' and f'' plots. Since the largest signal comes from the wavelength with maximal f'', the peak data set was collected at 0.9794 Å. An additional wavelength, which was 100 eVs away from the absorption edge, was selected as a remote wavelength (0.9611 Å). The remote data set was collected before the inflection data set because the crystal showed signs of radiation decay in *SCALEPACK* analysis (Otwinowski & Minor, 1997). The inflection wavelength was selected to have a maximal |f'| value and therefore an inflection data set was collected at 0.9795 Å. At Each wavelength, 360 ° phi rotation was collected in order to ensure the collection of Bijvoet pairs with great redundancy. Each frame was collected with either 8 or 10 sec exposure time and 0.4 ° - 0.6 ° oscillation ranges. The crystal to detector distance was either 200 or 250 mm. All the MAD experiments were done at beamline 17-ID operated by IMCA-CAT at the APS using ADSC Q210 detector.

All the three wavelength data sets were separately processed and scaled in MOSFLM / *SCALA* routines from the CCP4 suite of programs (Powell, 1999; Evans, 1997). Since the data sets at each wavelength were collected with high redundancy (~13 overall), the anomalous signals were kept separate during scaling and merging. To deduce the quality and phasing potential of SeMet derivative, **N**ormal **P**robability **A**nalysis of three wavelengths was performed using *FHSCALE* and *SCALEIT* programs in CCP4. *SCALEIT* analysis showed that the dispersive differences were smallest between peak and inflection (normal probability for acentric = 1.357) and largest between inflection and high energy remote (normal probability for acentric = 2.099). The data analysis also suggested that

anomalous differences (FpHi+ v. FPHi-) were highest for peak data set (normal probability for acentric = 1.934). The values for **N**ormal **P**robability **A**nalysis were well within the suggested range for a good derivative (Howell & Smith, 1992).

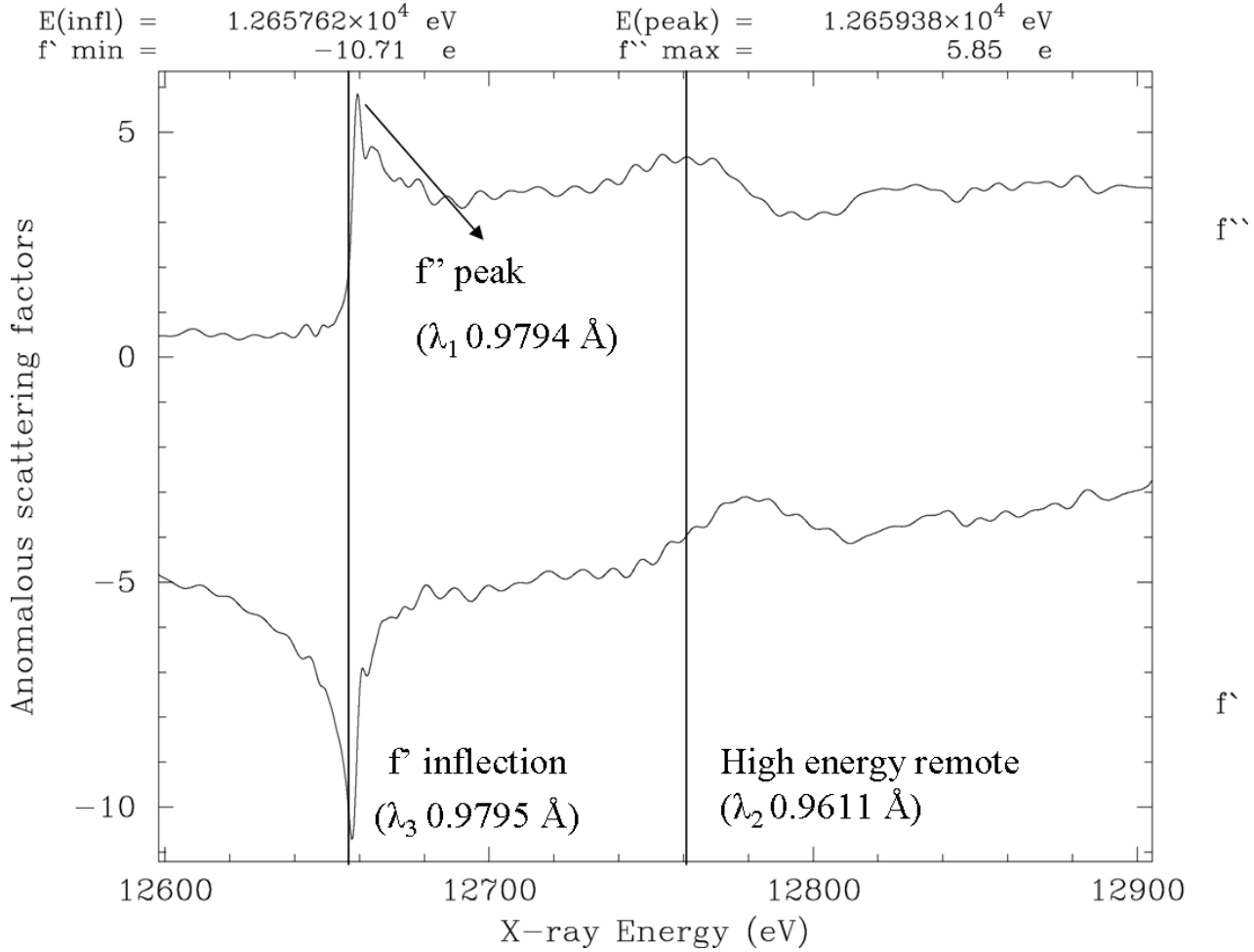


Figure 3.4 X-ray absorption spectrum of SeMet enriched ES* crystals.

3.1.8 Location and Refinement of 52 Selenium Sites

Data integration and scaling statistics of the SeMet-ES* crystal suggested an orthorhombic space group but intensity v/s *hkl* plots did not show a clear pattern of systematic absences. Therefore, the screw axis and space group could not be determined unambiguously. Space groups *P*222, $P222_1$, $P2_12_12$ $P2_12_12_1$ with *a*, *b*, *c* cell convention and other combinations with *a* having the screw axis, *b* having the screw axis, *a* and *b* having the screw axes, *b* and *c* having the screw axes, showed reasonable overall R_{syms} values (~10%). Space group ambiguity coupled with a large number of expected selenium sites (52) complicated the phasing problem. In order to deal with this combination of problems, a single wavelength peak data set was indexed in lowest symmetry space group (*P*222). The reason for selection of single wavelength peak data set in the beginning was to maximize the signal-to-noise ratio in order to have a reasonably large **ΔF** (difference in structure factor) value. In principle, the peak wavelength has the largest expected difference in $\mathbf{F}^+$ and $\mathbf{F}^-$, since the peak data set is collected at f" maxima. The choice of starting with a SAD data set also eliminated the problem of decreasing signal-to-noise ratio in subsequent data collection runs because of the radiation damage. The .mtz file obtained after indexing in *P*222 space group was reindexed in all possible space groups using a *CCP4* script. Each possible space group, was searched for Patterson cross peaks for selenium atoms. An automated program (*SOLVE* v. 2.03) was used to solve the patterson functions (Terwilliger & Eisenberg, 1983; Terwilliger *et al.*, 1987). Several *SOLVE* runs were started with different .mtz files between resolution range of 20 – 3 Å. The program gave the position of heavy atom sites, which were then evaluated based on Figure of Merit (FOM) and a statistical z score

(Terwilliger & Berendzen, 1997, 1999). *SOLVE* found a maximum number of selenium sites (25) in space group $P2_12_12_1$. The results in the space group $P2_12_12_1$ had the highest FOM = 0.43 and a z score of 169.53 in contrast to other space groups, which had low FOM ~ 20 and z scores ~ 25. Although a high FOM and z score value clearly pointed towards $P2_12_12_1$ space group, electron density maps constructed using the phase information from *SOLVE* were not interpretable. The density modification in *RESOLVE_huge* v. 2.06 (Terwilliger, 2001), which was explicitly compiled for large number of reflections, improved the FOM to 0.66 but did not give interpretable maps. The 25 selenium sites found from the SAD data analysis in *SOLVE* were input in *ADDSOLVE* (a routine in *SOLVE_giant* v. 2.06) to find additional selenium sites. Two separate *ADDSOLVE* runs were started at 3 Å resolutions. The first one used a 3 wavelength MAD data set scaled in *SCALEIT* and the second one used a 2 wavelength MAD (peak and remote) data set. Both the runs took more than a month to run on SGI (IRIX 6.5), which had 1GB swap space and dual 550 MHz processors. *ADDSOLVE* results with 3 wavelengths found a total of 30 selenium sites with the correct hand (FOM = 0.45 and z score = 251). These sites along with another 22 sites, which were found from analysis of residual maps in *SHARP* v. 3.0.15 (de LaFortelle & Bricogne, 1997) resulted in interpretable maps. The phases were then extended to the highest resolution shell (2.2 Å) using SHARP v.3.0.15. 5 % of the total reflections were taken out for cross validation of the model (Brünger, 1992).

3.1.9 Density Modification and Model Building

Since protein crystals typically contain 30-70% solvent, organized in channels of unordered water molecules, solvent flattening procedures like *DM* (Cowtan & Main, 1996), *SOLOMON* (Abrahams & Leslie, 1996) or *RESOLVE* (Terwilliger, 2002) often improve the quality of maps. In *SOLOMON* (CCP4), the electron density is constrained towards a flat solvent region, and this real-space density modification is iterated with a phase-combination step in reciprocal space. *DM* (CCP4) uses a similar iterative procedure in histogram matching where prior information in the form of expected density histograms is applied as constraints on the electron density map. *RESOLVE*, another density modification program, applies maximum-likelihood principle to do solvent flattening. The highest quality solvent flattened maps for SeMet-ES* to 2.2 Å were obtained from *SOLOMON* (Abrahams & Leslie, 1996) from CCP4. The FOM of the solvent flattened maps had a phenomenal value of 0.94 and the maps were of superb quality. Initial protein backbone tracing was done using the pattern recognition algorithm in *RESOLVE_huge* v.2.06 (Terwilliger, 2002). The remaining protein model was built manually in *O* using 1.5 σ level $2mF_o$ - DF_c electron density (Jones *et al.*, 1991) maps. At first, the model building in one chain was completed. The remaining three chains in the asymmetric unit were traced by applying the symmetry matrices generated in CCP4. Rotamers for the amino acid side chains were fitted using *O* library files. The N5-(2-nitrobutyl) modified FAD was manually modeled in the electron density using *O* dictionary, connectivity, torsion and stereo files generated from *PRODRG server* (Schuettelkopf & van Aalten, 2004).

3.1.10 Refinements (Background)

The initial protein model generally contains errors and must be optimized. The goal of crystallographic refinement can be formulated as finding the set of atomic coordinates that results in the best fit of the observed structure factor amplitudes and the amplitudes calculated from this model. In conventional least-squares refinement (Drenth, 1999), this goal has been formulated as finding the minimum of target function (3.12):

$$E_{X-ray} = \sum_{hkl} w_{hkl} \left(\left| F_{hkl}^{obs} \right| - k \left| F_{hkl}^{calc} \right| \right)^2 \qquad (3.12)$$

where w_{hkl} is a weighting factor, k is a scale factor and the calculated structure factor amplitude $|F^{calc}|$ is dependent on the parameter set of the model. Calculation of derivatives of $|F^{calc}|$ towards the model parameters allows application of gradient-driven optimization techniques to minimize this function. The refinement converges readily to the correct solution after many cycles if there are no false minima and the starting model is not drastically different from the actual structure (Sussman *et al.*, 1977).

However, the starting model is usually not very similar to the actual macromolecular structure. The reason for this is the poor observations (intensities) to parameters (three atomic positions x,y,z, one isotropic temperature factor B and occupancy for each non-hydrogen atom) ratio at the resolutions to which macromolecular crystals normally diffract (1.5 - 3.5 Å). To avoid over-fitting of the limited amount of experimental data, either more observations are added in the form of restraints, or the number of parameters are reduced by constraining the model in different ways (Sussman *et al.*, 1977). Typical

set of restraints come from the prior geometrical knowledge, which is expressed in the form of expected bond lengths, bond angles and van der Waals contact distances. Restraints are generally defined to conserve the planarity of planar groups or to maintain the chirality of chiral centers. The restraints are then entered as terms in the refinement target, which is defined as follows:

$$E = E_{geom} + wa\,E_{X-ray} \tag{3.13}$$

with E_{geom} being expressed as follows:

$$E_{geom} = \sum_{bonds} w_{bond}\,(r_{bond}^{ideal} - r_{bond}^{model})^2 + \sum_{angles} w_{angle}\,(\theta_{angle}^{ideal} - \theta_{angle}^{model})^2 + ... \tag{3.14}$$

where bond distances r_{bond}, angles θ_{angle}, and other geometric parameters of the protein model like torsion angles, planarity of rings *etc.* are restrained towards their ideal values using weights w. Weight *wa* (in equation 3.13) for the crystallographic term is chosen such that approximately equal gradient contributions from both sides of the combined target function result. Application of the constraints is only justified if X-ray diffraction data is of moderate resolution (d_{min} 2.5 Å). The number of parameters is reduced by either constraining certain groups of atoms or even an entire molecule to a rigid body. This limits the degrees of freedom to torsion angles, thus resulting in a further improved observation-to-parameter ratio. The rigid body refinement might bring a model within a radius of convergence for least-squares. A torsion-angle parameterization of protein

molecules has been implemented in the *CNS* program (Brünger *et al.*, 1998a; Rice & Brünger, 1994).

Molecular dynamics simulations allow greater energy barriers to be overcome in the positional parameter variation which helps to avoid the refinement becoming trapped in local minima. The technique simulates possible energetically allowed structures for a given temperature and pressure. The process is repeated after a short time step with the atoms in the new positions until a minimum potential energy is reached describing the dynamic behavior of the atoms. The temperature can be raised in order to overcome energy barriers and then decreased slowly to approach the energy minimum (a process called simulated annealing (Kirkpatrick *et al.*, 1983). The temperature during the simulated annealing refinement is a parameter describing the height of the local energy barriers that the refinement can overcome. Hence, simulated annealing refinement has a significantly larger radius of convergence than the restrained least-squares refinement. Movements greater than 5 Å can be achieved (Brünger *et al.*, 1990; Rice & Brünger, 1994).

Major advances in refinement have been made by the formulation of maximum likelihood target functions (Bricogne, 1997). In contrast to least-squares methods, maximum likelihood provides a statistically valid way to deal with errors and incompleteness of the model. A general approach is to represent the resolution-dependent quality of the model by the σ_A distribution (3.15):

$$\sigma_A \quad = \quad \left\langle E^{obs} \; . \; E^{calc} \right\rangle \tag{3.15}$$

where σ_A values are calculated in resolution bins and E^{obs} and E^{calc} are observed and calculated normalized structure factors. Since the phases of E^{obs} are unknown, σ_A values need to be estimated. Cross-validation, initially introduced to monitor over-fitting of the data by calculation of a free R-factor (Brünger, 1997), plays an important role in the estimation of σ_A values (Adams *et al.*, 1999). In cross-validation, typically 5-10% of the data (the test set) is kept outside the refinement. Estimation of σ_A values based on these test set reflections avoids serious over estimation resulting from over-fitting of the data. The probability to observe E^{obs}, given E^{calc} of the model, can then be calculated by (3.16):

$$P(E^{obs}\,;\,E^{calc}) \quad = \quad \frac{1}{\pi\,(1-\sigma_A^2)} \;\; \exp\left(\; - \; \frac{\left| E^{obs} - \sigma_A \, E^{calc} \right|^2}{1 \quad - \sigma_A^2} \right) \tag{3.16}$$

Similar equations are derived to calculate the probability to observe structure factor amplitude $|F^{obs}|$, given calculated structure factor F^{calc} and the measurement error in $|F^{obs}|$. Maximum likelihood refinement aims to maximize the likelihood of measuring the set of observed structure factor amplitudes, given the calculated structure factors of the model.

A recent development in protein structure refinement is the possibility to model anisotropic motions of complete domains by TLS-parameterization for the translation, libration and screw-rotation displacements of pseudo-rigid bodies (Schomaker & Trueblood, 1968) as implemented in the maximum likelihood refinement program

*REFMAC*5 (Winn *et al.*, 2001). Despite the developments mentioned above, the radius of convergence of protein structure refinement remains limited and in current practice refinement cycles still need to be iterated with time-consuming rebuilding steps where the model is improved manually by interpretation of electron density maps.

3.1.11 Refinements of ES* Crystal Structure

After building significant portions of the ES* homotetramer in the asymmetric unit, rigid body refinements were performed between resolution range of 40 – 2.2 Å. All the residues of the homotetramer in the asymmetric unit were specified as one rigid body. Molecular topology file for the refinements was generated using generate_easy routine in *CNS*. Initial refinements were carried out against the merged F_{obs} from all the three MAD wavelengths. A maximum likelihood target including amplitudes and phase probability distribution (Brünger *et al.*, 1998b) was chosen in initial refinements. Twenty minimization steps of rigid body refinement were performed along with initial isotropic B-factor refinement and bulk solvent correction. At this stage of the refinements, the N5-FAD adduct was not modeled into the electron density maps. After rigid body refinements R_{cryst} (Brünger, 1997) dropped to 46.68 % from an initial value of 52.84 % whereas R_{free}(Brünger, 1992) (calculated from the 5 % of the total reflections randomly selected from the entire MAD data set) dropped to 46.55 % from a starting value of 52.72 %. Since most of the model fitted well into the electron density ($2mF_o$ - DF_c) maps, simulated annealing runs were performed only with torsional molecular dynamics (Rice & Brünger, 1994). The Cartesian molecular dynamics was constrained in all the four cycles of simulated annealing. The starting protein model was heated to high temperature

of 10,000 K for annealing procedure with a step size of 25 K. The model was slow cooled to a final temperature of 100 K. The R_{cryst} after the final run of simulated annealing dropped from 44.92 % to 28.06 % whereas R_{free} dropped from 44.21 % to 31.56 %. After simulated annealing, B factors were refined by choosing the whole homotetramer as one group. Individual B factors were also refined afterwards. The B factor refinements dropped the R_{cryst} to 25.31 % and R_{free} to 28.9 %.

Unambiguous electron density for N5-(2-nitrobutyl) moiety was observed at 5 σ level in mF_o - DF_c maps. The N5 modified adduct was modeled using parameter and topology files generated from *PRODRG server* (Schuettelkopf & van Aalten, 2004). Both *cis* and *trans* form of the nitrobutyl moiety were modeled into the electron density in separate trials. The refinements of *cis* model showed negative electron density at 3 σ on the 4th carbon atom. The *trans* form showed no negative density in mF_o - DF_c maps C4 atom of the nitrobutyl moiety was replaced by electron dense nitro group. The *cis* form of the nitrobutyl moiety was also ruled out because the nitro group in this conformation was too close to side chain oxygen of the Ser171 and O4 of the isoalloxazine ring. Although refinements were converging as evident from the decrease in the R_{cryst} values, all the seleniums showed negative electron density at 3 σ levels in mF_o - DF_c maps. The problem was partially solved by using anomalous libraries (generated in *CNS* using experimental f' and f'' values from the fluorescence scans of SeMet) and refining the calculated structure factors against the observed structure factors from the peak data. Apparently, the selenium signal was bleached over the period of data collection and the intensity of selenium signal was proportional to the wavelength at which the data was

collected. Out of 52 selenium sites, 24 selenium sites still had negative density at 3 σ level but had very low B factors (~28 $Å^2$). In order to account for the negative density on selenium atoms, B factor restraints were relaxed to the following values:

1) Main chain bonds = 2.0 $Å^2$ (default = 1.5 $Å^2$)

2) Main chain angles = 3.0 $Å^2$ (default = 2.0 $Å^2$)

3) Side chain bonds = 4.5 $Å^2$ (default = 2.0 $Å^2$)

4) Side chain angles = 6.0 $Å^2$ (default = 2.5 $Å^2$)

Finally, the occupancy of selenium sites was refined using q_group routine in *CNS* v. 1.1. In later stages of refinements selenium occupancies were manually optimized by looking at the negative density features, which were still present in the mF_o - DF_c maps at 3 σ level.

A total of 16 cycles of refine.inp were run in order to achieve a final R_{cryst} value of 23.68% and R_{free} of 26.72 %. A maximum likelihood function using amplitudes was chosen as target in the later stages of refinements. All the refinements were done using *CNS* v. 1.1. After the refinements had converged water molecules were added to the protein homotetramer using *ARP/wARP* (Perrakis *et al.*, 1999) routine in *CCP4*. In first cycle a total of 400 water molecules were added with a sigma cut off of 4. In the second cycle, another 600 water molecules were added with *ARP/wARP* with sigma cut off of 2.5. Water molecules with B factor greater than 45 $Å^2$ were inspected manually using *O*. The difference features in the mF_o - DF_c maps overlapped with $2mF_o$ - DF_c maps were manually inspected to add or delete more water molecules. The final structure was refined with one homotetramer (1720 protein residues), 4 N5-FAD adducts and 694 water

molecules per asymmetric unit. The structure of ES* was refined with final R_{factor} of 20.16 % and R_{free} of 23.22 % to 2.2 Å resolution. Secondary structure analysis was done using *KSDSSP* (Kabsch & Sander, 1983).

3.2 Crystal Structure Determination of Oxidized Form of NAO

3.2.1 Preparation and Purification of Natural and SeMet Enriched Form of Oxidized NAO

Recombinant nitroalkane oxidase was expressed and purified from *E. coli* strain BL21 (DE3) transformed with pETNAO4 as previously described (Daubner *et al.*, 2002; Valley & Fitzpatrick, 2003b). The purified enzyme was stored in 20 mM HEPES pH 8.0, 1 mM EDTA, 5% glycerol and aliquots were flash frozen with liquid N_2. Se-Met enriched NAO was obtained by expression in the methionine auxotroph *E. coli* B834 (DE3) (Novagen) (Nagpal *et al.*, 2004). A single colony of *E. coli* B834 (DE3) transformed with pETNAO4 was used to inoculate 5 ml of LB containing 100 µg/ml carbenicillin at 37 °C. After 8 hours, 1 ml of the starter culture was used to inoculate 50 ml of a minimal media composed of the following (g/l except where noted): Na2HPO_4, 7.0; KH2PO_4, 3.0; NH_4Cl, 1.0; all amino acids, 0.1; adenine, 0.1; thiamine and riboflavin, 0.002; carbenicillin, 0.1; $MgSO_4$, 2 mM; $CaCl_2$, 0.1 mM; and D-glucose, 22 mM. After 12 hours at 37 °C, 4 ml of the step up culture was used to inoculate 1 liter (6X) of the same minimal media, except methionine was replaced with seleniomethionine at 0.1 g/l. When the absorbance at 600 nm reached 0.6, the temperature of the cultures was lowered

from 37 to 25 °C and IPTG was added to a final concentration of 0.25 mM. Cells were harvested after 8.5 hours of expression and Se-Met enriched NAO was purified by the normal protocol. The gene sequence for NAO predicts a 439 amino acid protein of 48162 Dalton. The amino terminal methionine is cleaved by the expression host resulting in 13 methionines per chain. Biophysical solution studies are consistent with a homotetramer-homodimer equilibrium with a Ka for tetramerization of 8 x 10^{6} M^{-1} (Gadda & Fitzpatrick, 1998).

3.2.2 Crystallization of Oxidized NAO in Three Different Space Groups

Needle shaped microcrystals were initially obtained from cryo screen HR-1 (Hampton Research, Aliso Viejo, CA) with the hanging drop, vapor diffusion method. These conditions were optimized and diffraction quality crystals **(form 1)** were grown by mixing 2 µl of 10 mg/ml protein solution (10 mM sodium cacodylate, pH 7.5) with 2 µl of reservoir solution containing 25% w/v PEG 4000 and 35% v/v glycerol, sodium cacodylate trihydrate (200 mM, pH 7.5), 1 mM spermine hydrochloride (Spm-Cl) at 4 °C (Figure 3.5 *a*). The crystals typically grew within 10 d. NAO in crystal **form 2** was grown from conditions containing 25% w/v PEG 4000 and 30% v/v glycerol, sodium cacodylate trihydrate (200 mM, pH 7.5), 1mM Spm-Cl and 8% w/v of 1,6-hexanediol at 4 °C. Single, square shaped crystals (0.15 x 0.12 mm) were obtained by mixing 2 µl of 8 mg/ml of protein with 2 µl of reservoir solution after 25-30 d equilibration at 4 °C (Figure 3.5 *d*). Crystal **form 3** was from SeMet-NAO enzyme grown from conditions similar to crystal form 1 except the protein was incubated at 4 °C for 5 h with 10 mM dithiothreitol prior to setting up the drop (Figure 3.5 *g*). The SeMet-NAO crystals

typically grew within 2 weeks. For each crystal form, single crystals were mounted directly in nylon loops since no additional cryoprotectant was necessary, and they were flash frozen by quick submersion into liquid N_2 (Nagpal *et al.*, 2004).

3.2.3 Data Collection & Scaling of Oxidized NAO Crystals (Form II & III)

X-ray diffraction data from crystal form II were collected at the SER-CAT sector of APS on beamline 22-ID using a marCCD 165 detector. Each image was collected using 1.0 Å x-rays with a crystal to detector distance of 250 mm. A total of 180° of phi rotation was collected with a 2 sec exposure time per frame and 0.5º oscillation range (Figure 3.5 *e, f*). A three wavelength MAD dataset on SeMet-NAO was collected at SER-CAT using a mar225 detector. A high sensitivity fluorimeter (Röntec USA Inc., Carlisle, MA) was used to separate the arsenate peak from the selenium signal. Forward and inverse beam datasets (110° of 30 phi rotation, each) were collected at the peak (0.9792 Å), inflection (0.9794 Å) and remote (0.9686 Å) wavelengths. Each frame was collected with an exposure time of 4 s and 1.0° oscillation range (Figure 3.5 h, i). The crystal to detector distance was 270 mm (Nagpal *et al.*, 2004). Datasets for crystal form III were processed with *HKL-2000* (Otwinowski & Minor, 1997). The datasets for crystal form II were processed with *MOSFLM* (Powell, 1999) and *SCALA* (Hamilton *et al.*, 1965). The three wavelength MAD dataset of SeMet-NAO was further scaled and analyzed with *FHSCALE* and *SCALEIT* from the *CCP4* suite of programs (Howell & Smith, 1992).

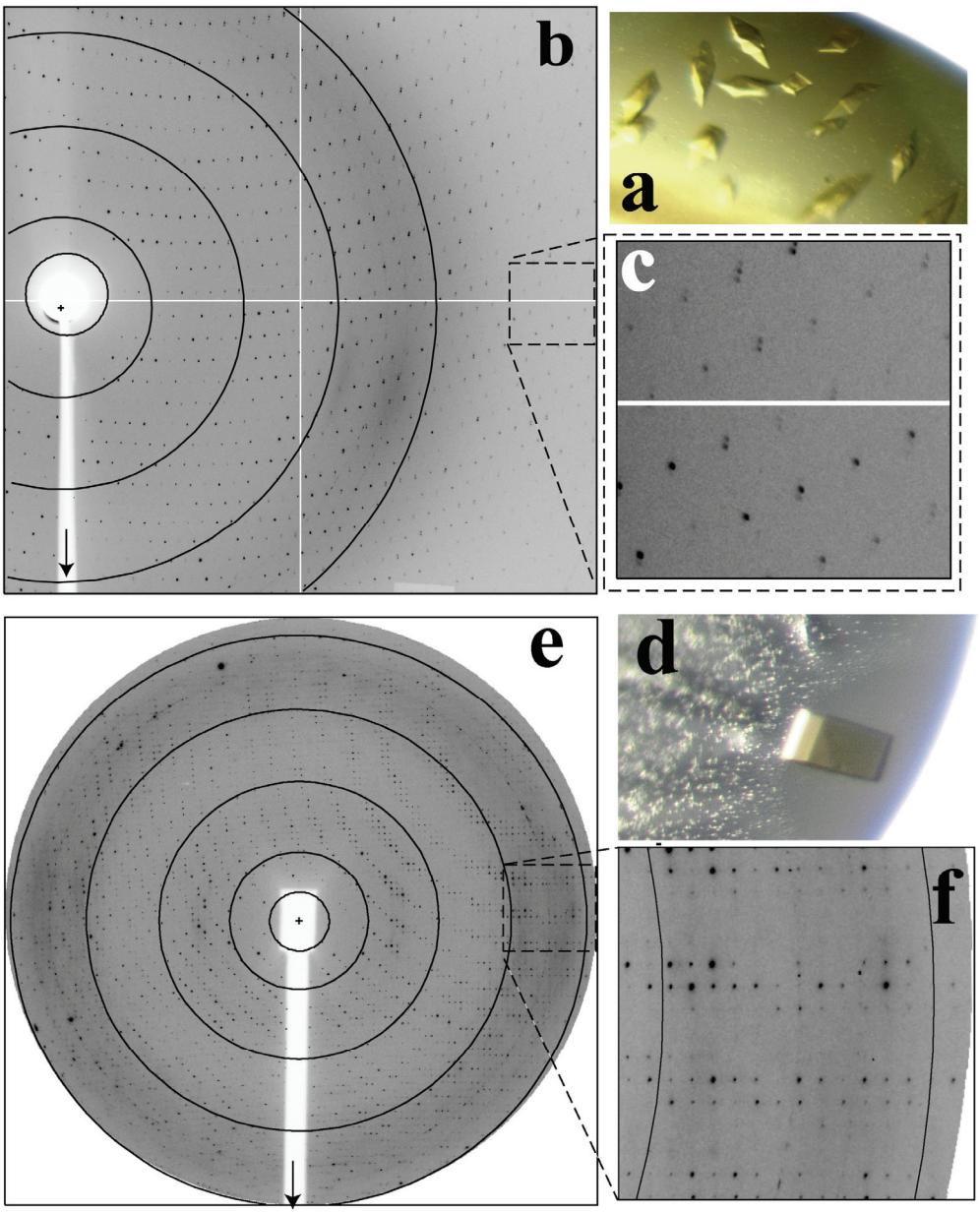


Figure 3.5 Crystals photographed without polarization of nitroalkane oxidase (NAO) in a) crystal form 1 (0.25 X 0.15 mm), d) crystal form 2 (0.15 X 0.12 mm). The x-ray diffraction pattern at 100K with rotation about the vertical axis (arrows) collected with b) ADSC Quantum 4 detector at BIOCARS beamline 14-BMC from (a) with 60 s exposure, 0.3º phi rotation, 400 mm crystal-to-detector, arcs indicate (3.5, 4.6, 6.9, 13.8 Å), 2θ detector offset = 10.7º; e) from (d) but collected with a marccd 165 detector at SER-CAT beamline 22-ID, 2 s exposure, 0.5º phi rotation, 250 mm crystal-to-detector distance, circles indicate (3.4, 4.5, 6.8 and 13.6 Å). An expanded and contrast adjusted view of the high resolution diffraction perpendicular to the rotation axis for: c) from (b), 2.5 Å at the edge, f) from (e), 3.4 Å at the edge.

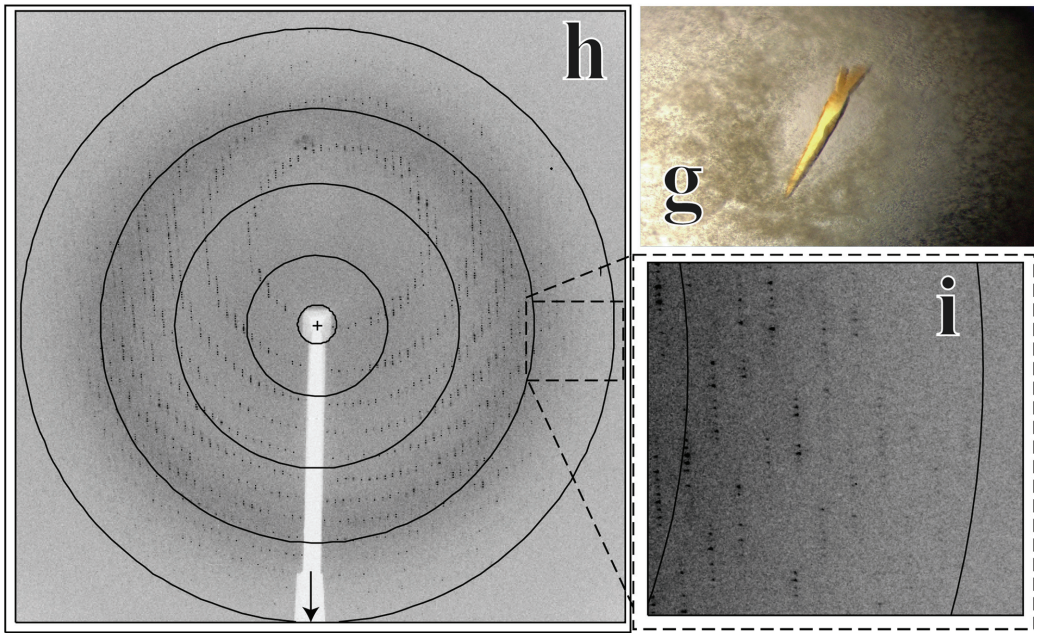


Figure 3.5 Continued g) SeMet enriched crystals of oxidized NAO photographed without polarization in crystal form 3 (0.3 X 0.1 mm). The x-ray diffraction pattern was collected at 100K with rotation about the vertical axis (arrows). h) was collected from (g) with 4 s exposure collected with a marccd 225 detector at SER-CAT, 1.0º phi rotation, 270 mm crystal-to-detector distance, circles indicate (2.4, 3.3, 4.9 and 9.8 Å). An expanded and contrast adjusted view of the high resolution diffraction perpendicular to the rotation axis form (h) is shown in (i). The resolution at the edge of the detector was 2.8 Å.

3.2.4 Data Collection and Processing Strategy for 485 Å Long Cell Edge Crystals of Oxidized NAO (Form I)

[§]Oxidized NAO crystallized easily in form I and often diffracted to very high resolution (1.6 Å), which encouraged us to vigorously pursue this crystal form. However, the diffraction pattern observed with our in-house Cu Kα X-rays was very streaky, even

[§] Reformatted from http://www.nsls.bnl.gov/newsroom/science/2005/03-Orville.htm (Authors: Nagpal, A. and Orville, A.M.)

at the maximum crystal to film distance (Figure 3.6). This prompted us to carefully alter our crystallization conditions, cryo-conditions, freezing processes and general manipulation techniques. However, the diffraction patterns were always disappointing relative to the visual qualities of the crystals. In contrast, our first diffraction images obtained at X26C gave different results (Figure 3.6). The NAO crystals diffracted to beyond 1.6 Å resolution in space group $P3_221$, but with unit cell dimensions of $a = b$ =104 Å, c = 485 Å. Since the observed X-ray diffraction pattern has a reciprocal relationship with the unit cell dimensions of the crystal in real space (Figure 3.7), the X-ray reflections collected from NAO long cell edged crystals were always closely spaced. The close diffraction spots could not be resolved at home source (Rigaku IV++ X-ray generator) because of the long wavelength (CuKα 1.54 Å) and the limitations of the beam optics (Figure 3.6). In contrast, better beamline optics, shorter wavelength (0.9 Å) and good point resolution detectors at synchrotron beamlines resolved the diffraction spots. Along with the above mentioned parameters, large crystal to detector distance was also required for resolving the closely spaced spots. The large distance allowed the X-ray reflections to diverge from each other before they reached the detector. However, this also significantly decreased the maximum Bragg angle of reflections that reached the detector and thus the high resolution reflections was not collected. Because of this reciprocal relationship (Figure 3.7), the complete data set to high resolution could not be easily achieved. In order to collect a complete data set to high resolution, we visited several beamlines at APS and NSLS and tried different area detectors (mar165, ADSC Quantum 4, ADSC Q210, mar225, Bruker Proteum 300). Although area detectors like Bruker Proteum 300 had the maximum active area (300 X 300 mm) for data collection,

the point resolution function of the detector was not good enough to resolve the spots. Other data collection parameters like 2θ offsets and κ geometry were also optimized during different data collection runs.

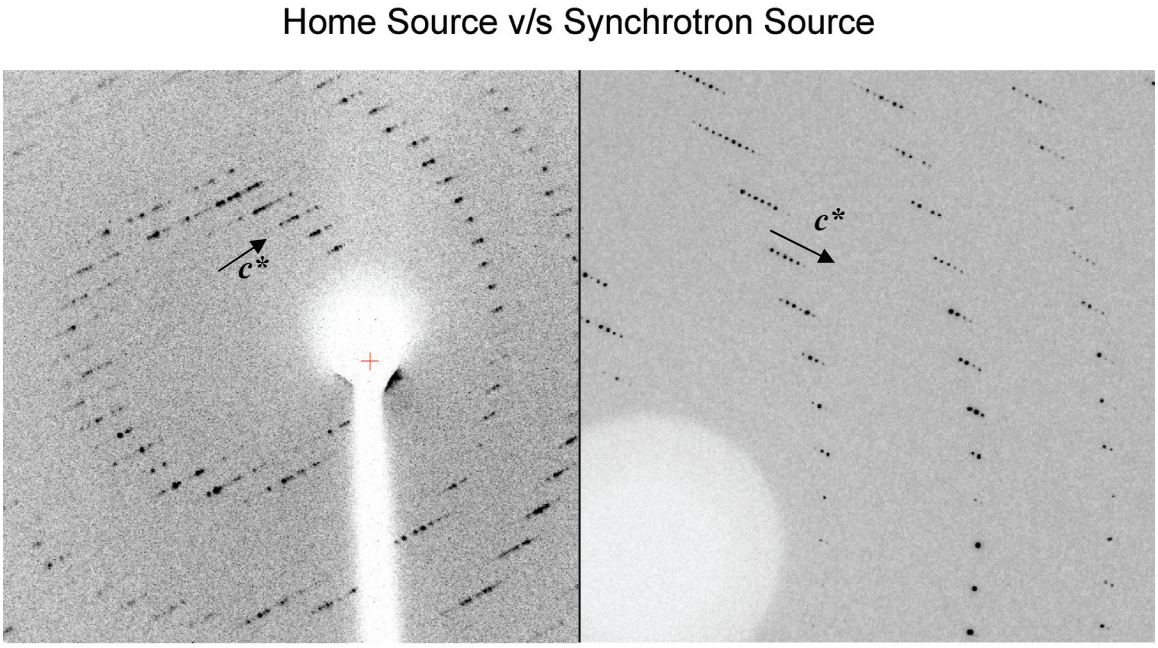


Figure 3.6 Expanded sections of the observed X-ray diffraction pattern collected from long cell edge NAO crystals with a detector distance of 400 mm. The data on the left was collected with Raxis IV++ X-ray generator at home where as the data on the right was collected at a synchrotron source.

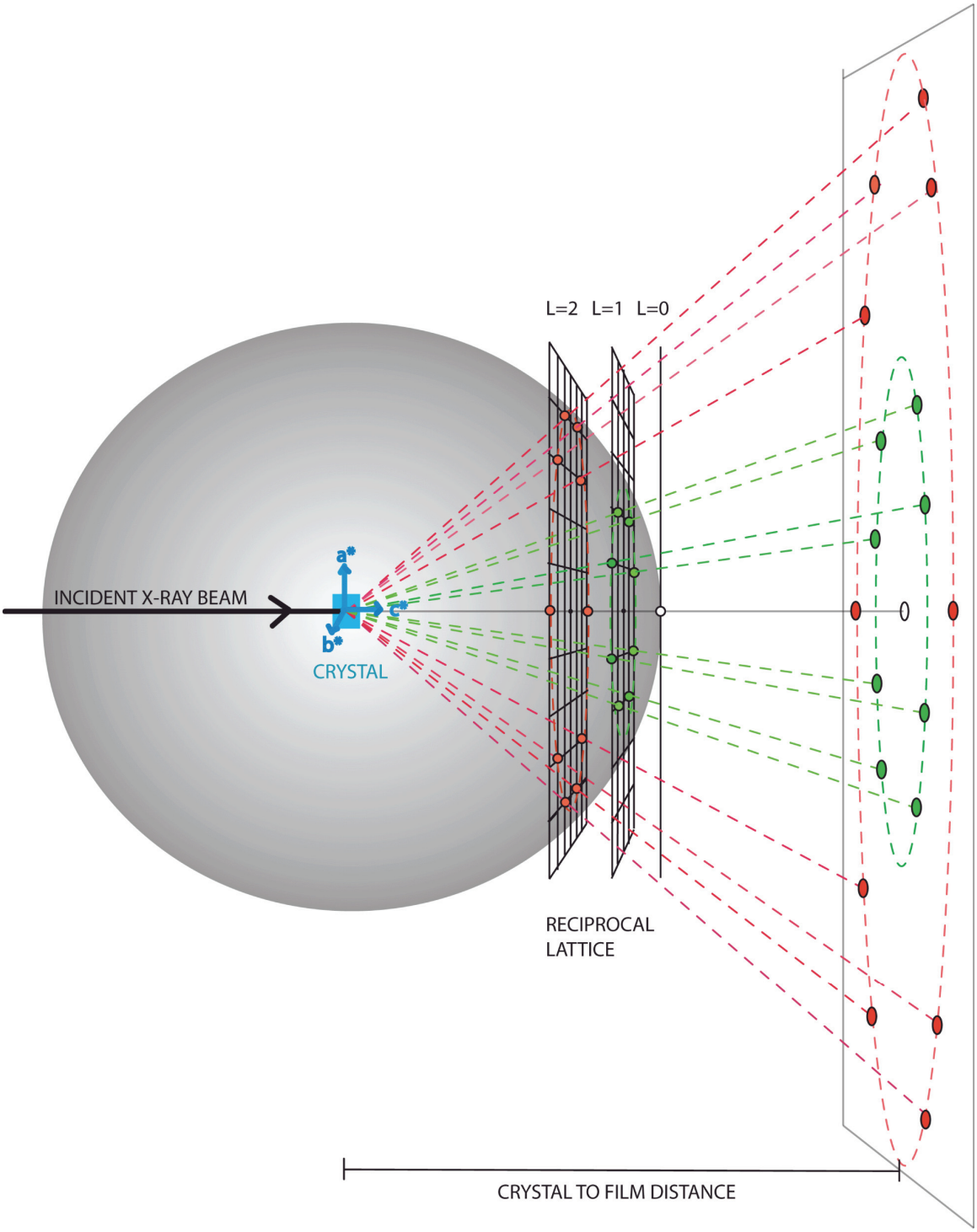


Figure 3.7 Ewald sphere construction showing the reciprocal relationship between real space unit cell and the observed diffraction pattern[**]

[**] reproduced from http://www.doe-mbi.ucla.edu/%7Esawaya/m230d/

For the high resolution data of NAO, our strategy was to collect several data sets that differed with a vertical offsets (approximately equivalent to 2θ and to optimize the kappa geometry such that the long cell axis was approximately parallel to the crystal rotation axis and to fill in missing regions of reciprocal space. The data collection strategy, which finally led to structure solution of oxidized form of NAO consisted of three or four data collection sweeps using 0.90 Å X-rays and an ADSC Quantum 4 detector at the BioCARS beamline 14-BMC of the Advanced Photon Source (APS) at Argonne National Laboratory. First, the high resolution data was collected at κ = 0° and a total of 120° of phi rotation with 0.3° rotation per frame. The detector was vertically offset 150 mm and the crystal to detector distance was 400 mm. Next, kappa was rotated to ± 40° and omega sweeps of 100° total rotation were collected with the same vertical offset and detector distance. The detector was then moved to 80 mm vertical offset and 120° of phi rotation were collected with all of the other parameters as in the second pass. Finally, the low resolution data was collected with 100° of phi rotation in 0.3° rotation per frame with κ = 0°, no vertical detector offset and a distance of 400 mm. Each frame was collected with an exposure time of 30 sec. Data sets were processed in *HKL2000* and merged in *SCALEPACK*. Merging these data sets yielded overall 93.5% completeness and 80% completeness in the 2.07 Å resolution shell, but only 60% complete to 1.6 Å resolution. As of June 2005 only 172 x-ray structures (out of 314,34) deposited in the PDB contain a *c* axis unit cell edge greater than 480 Å. Moreover, only 17 of these structures were to at least 2.5 Å resolution. To our knowledge, the structure of oxidized form of NAO sets the highest resolution record for a large unit cell in the Protein Data Bank (PDB) (Nagpal *et al.*, 2004).

3.2.5 Molecular Replacement Solution for Oxidized Form of NAO

The data processing of the crystals in form I showed a pattern of systematic absences. Every third reflection was present (special condition l = 3n), which narrowed the space group choices to $P3_1$, $P3_2$, $P3_112$, $P3_212$, $P3_121$, $P3_221$, $P6_2$, $P6_4$, $P6_222$, $P6_422$. Out of these possibilities, reasonable R_{sym} (~7.6 %) were only observed for $P3_1$, $P3_2$, $P3_121$, $P3_221$ space groups. The possibility of the other trigonal and hexagonal space groups ($l \neq 3n$) was not totally eliminated because the observed pattern of systematic absences could also arise from non-crystallographic symmetry. Molecular replacement efforts were started in $P3_1$, $P3_2$, $P3_121$, $P3_221$ space groups. Because of the large unit cell Matthews co-efficient calculations (Kantardjieff & Bernhard, 2003) were reasonable for 4 – 12 monomers per asymmetric unit. The statistical probability graphs for the solvent content calculations at 2 Å resolution suggested 4 – 8 molecules per asymmetric unit depending upon the space group chosen. A large number of molecular replacement runs were set up in *MOLREP* (*CCP4*) (Vagin & Teplyakov, 1997) to find 4 - 8 molecules per as asymmetric unit with ES* tetramer or monomer as search model. For preparing the search model, all the SeMet were changed to Met and FAD adduct and water molecules were removed. All the molecular replacement results were refined in order to evaluate the quality of the solution. 5% of the total reflections were isolated to cross validate the refined model (Brünger, 1992). The rigid body and restrained refinements were performed in *REFMAC5* (Murshudov *et al.*, 1997) using low resolution (3 Å) data. Results in all the space groups with varying number of monomers per asymmetric unit except for $P3_221$ with 6 monomers per asymmetric unit gave an R_{cryst} value varying between (52.6 % - 36. 2%) and R_{free} between the range of (57.1 % - 42.5%). The best

molecular replacement solution was obtained using ES* monomer as a search model in $P3_221$ with 6 monomers per asymmetric unit (correlation co-efficient – 62.3%). Rigid body and restrained refinement gave an R_{cryst} of 21.5 % and R_{free} of 26.6 %.

3.2.6 Model Building and Refinement Procedures for Oxidized Form of NAO

The molecular replacement solution matched the experimental electron density maps ($2mF_o$ - DF_c) and no significant model rebuilding was required. A total of 19858 protein atoms were manually inspected in the model using *O* (Jones *et al.*, 1991), which took a significant amount of time. Side chains and rotamers were adjusted using *O* library files. The FAD was modeled and refined using the *O* and *REFMAC5* dictionary files created with *PRODRG server* (Schuettelkopf & van Aalten, 2004). Maximum likelihood refinements were carried against the F_{obs} calculated between resolution of 50 – 2.07 Å resolution using *REFMAC5* in *CCP4*. A long tubular positive density at 3.5 σ (mF_o - DF_c maps) was observed in the active site of the three out of the six subunits in the asymmetric unit. Spermine, a weak competitive inhibitor of NAO ((K_i ~200 mM at pH 8), which was present in the crystallization conditions, was fitted into the electron density using the *O* dictionary files from the *HIC-Up server* (Kleywegt & Jones, 1998). The electron density for spermine in the three subunits was refined in *REFMAC5* with an overall B factor of 40 $Å^2$. After the refinements had converged, water molecules were added using *ARP/wARP* (Perrakis *et al.*, 1999) routine in *CCP4*. In first cycle a total of 500 water molecules were added with a sigma cut off of 4. In the second cycle, another 800 water molecules were added with *ARP/wARP* with sigma cut off of 2.5. Water

molecules with a B factor greater than 45 $Å^2$ were manually inspected using *O*. The difference features in the mF_o - DF_c maps overlapped with $2mF_o$ - DF_c maps were manually inspected to add/ remove water molecules, glycerols and PEG4000. The final refined model had 1.5 homotetramer (2580 protein residues), 1136 water molecules, 14 glycerol molecules and 2 PEG4000 molecules per asymmetric unit. After adding all the ligands and water molecules a positive density feature to 3.5 σ level (mF_o - DF_c maps) became apparent in the active site of the C chain. A spermine molecule was modeled into the active site of C chain. Spermine molecules were refined to an overall average B factor of 46 $Å^2$. The final structure of the oxidized NAO has two empty (E^{ox}) and four inhibitor occupied active sites (EI) with a final R_{factor} of 18.76 % and R_{free} of 22.51 % to 2.07 Å resolution.

4 RESULTS

This chapter describes data collection and phasing statistics of ES* and EI crystals. The statistical calculations were performed to assess the quality of X-ray diffraction data. Phasing statistics like *Phasing Power* and R_{cullis} were used to estimate the phasing potential of the derivative. Initially, heavy atom soaks with Pt and Au salts were used to derivatize ES* crystals. Although MIR (multiple isomorphous replacement) data sets had reasonable R_{syms} (~8%) and R_{iso} (the parameter used to describe the isomorphous differences between native and derivative crystals) values, the location of the heavy atom sites did not yield interpretable maps. The phasing power for MIR data sets was less than one, which led to the conclusion that heavy atoms in ES* derivatives were not localized and therefore could resolve the phase problem. Since MIR methods were not successful this chapter does not describe the data collection and phasing statistics for MIR data sets but only concentrates on the results from SeMet MAD data analysis. The quality of ES* and EI crystal structures was evaluated using Ramachandran Plots, temperature factor analysis and ESU (estimated standard uncertainity) values calculated form *REFMAC5* (Cruickshank, 1999). The three dimensional structures of ES* and EI were compared by calculating the r.m.s deviation plots using *SUPERPOSE* in *CCP4* (Krissinel & Henrick, 2004). A calculation of solvent accessible surface areas for all the residues in ES*, EI and E^{ox} structures was done using *CNS* v. 1.1 (Brünger *et al.*, 1998a). The surface accessibility of FAD and its atoms in different crystal structures of NAO was calculated using *EDPDB* (Zhang & Matthews, 1995).

4.1 Data Collection Statistics for ES* and SeMet-ES* Crystals

The ES* form of NAO was successfully crystallized in both natural abundance and SeMet enriched forms, as described in experimental section. The crystals were colorless suggesting the presence of the 5-(2-nitrobutyl)-1,5-dihydro-FAD adduct. To prevent conversion to active enzyme, all the crystal manipulations were done at 4 °C and all the diffraction data was collected from cryo-preserved crystals. A complete data native data set was collected to 2.5 Å resolution. Unit cell parameters and data collection statistics are presented in (Table 4.1). A very complete and highly redundant three wavelength MAD dataset was collected and analyzed for anomalous scattering from the selenium atoms in SeMet-ES* to 2.2 Å resolution (Table 4.1). Cell parameters were very similar to the native data set. Solvent content and Matthews coefficient analysis suggests the presence of one homotetramer per asymmetric unit (V_M 3.35; 63.3% solvent) (Kantardjieff & Rupp, 2003; Matthews, 1968). Self rotation function analysis indicates the presence of 222 fold local symmetry, which is consistent with four subunits in the asymmetric unit (Tong & Rossmann, 1997).

Table 4.1 Data collection statistics for ES* and SeMet ES*

Data Set	SeMet-ES* (Peak)	SeMet-ES* (Remote)	SeMet-ES* (Inflection)	Native ES*
X-ray Source	APS	APS	APS	NSLS
Beamline	IMCA-CAT 17-ID	IMCA-CAT 17-ID	IMCA-CAT 17-ID	X26C
Wavelength (Å)	0.9794	0.9611	0.9795	1.0007
Detector	ADSC Q210	ADSC Q210	ADSC Q210	ADSC Q4
Distance (mm)	200	200	250	200
Exposure time (s)	8	8	10	60
Resolution Range (Å) [a]	40 – 2.2 (2.32 – 2.2)	40 – 2.64 (2.78 – 2.64)	40 – 2.85 (3.00 – 2.85)	50 – 2.3 (2.42 – 2.30)
Δ phi per image (°)	0.4	0.4	0.6	0.5
Mosaic Spread (°)	~ 0.55	~ 0.55	~ 0.56	~ 0.3
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit Cell (Å), *a*	90.4	90.4	90.5	90.5
b	163.9	163.2	163.9	164.0
c	173.4	173.5	173.8	172.5
Total Reflections	1496015	1023723	850177	789148
Unique Reflections	130755	76209	61184	114527
Multiplicity	11.4	13.4	13.9	6.9
Completeness (%) [a]	99.9 (99.9)	99.9 (99.9)	99.9 (99.9)	99.9 (99.9)
Anomalous comp. (%)	98.9	99.8	99.8	-
R_{sym} [a,b]	0.105 (0.53)	0.106 (0.422)	0.097 (0.491)	0.133 (0.437)
I/σ(I) [a,c]	6.0 (1.5)	5.4 (1.7)	6.8 (1.5)	5.0 (1.6)

[a] Values for the highest resolution shell of data are given in parentheses. [b] R_{sym} (I) gives the average agreement between the independently measured intensities such as $\Sigma_h\Sigma_i |I_i - I| / \Sigma_h\Sigma_i I$, where I is the mean intensity of the i observations of reflection h. [c] I/σ(I) is the root-mean-square value of the intensity measurements divided by their estimated standard deviation.

4.2 Phasing Statistics for SeMet-ES* Crystals

ADDSOLVE subroutine from *SOLVE_giant* v. 2.06 (Terwilliger & Eisenberg, 1983; Terwilliger *et al.*, 1987) was used to locate 30 SeMet sites in $P2_12_12_1$ space group. The phasing results had an overall figure of merit (FOM = 0.45) and z score of 251. An additional 22 SeMet sites were located with *SHARP* (de LaFortelle & Bricogne, 1997) by analyzing residual maps based upon the 30 SeMet sites. All 52 SeMet sites refined to convergence to 2.2 Å resolution in *SHARP* (Table 4.2). Solvent flattening in *SOLOMON* (Abrahams & Leslie, 1996) produced clear and interpretable electron density maps (FOM 0.94).

Table 4.2 Phasing statistics for SeMet-ES*

	peak (f")	remote	inflection (f')
Wavelength (Å)	0.9794	0.9611	0.9795
Resolution (Å)	40 – 2.2	40 – 2.5	40 – 2.8
R_{cullis} centric (iso)[b]	0	0.99	0.99
R_{cullis} accentric (iso/ano)	0 / 0.758	0.978 / 0.840	1.0 / 0.839
Phasing power centric (iso)[c]	0	0.351	0.138
Phasing power (iso/ano)	0 / 1.405	0.361 / 0.974	0.136 / 0.80
FOM[d]	0.13 (centric reflections) 7950 0.32 (accentric reflections) 119500 0.94 after solvent flattening in SOLOMON		

[a] Compiled from SHARP (de LaFortelle & Bricogne, 1997).
[b] R_{cullis} = <phase-integrated lack of closure> / < | F_{ph} - F_p | >, where F_{ph} is the structure factor obtained from the inflection, peak data sets and F_p is obtained from the remote data set. The isomorphous and anomalous differences are designated as iso and ano, respectively.
[c] Phasing power = < [| $F_{h(calc)}$ | / phase-integrated lack of closure] >, where $F_{h(calc)}$ is the calculated structure factor for 52 Se atoms.
[d] Figure of merit statistics describe a confidence level for the calculated centroid phases.

4.3 Evaluation of the Model Quality (ES*)

The final model consists of one homotetramer per asymmetric unit. A total of 1720 residues were built in the asymmetric unit, 430 per subunit. The electron density for eight C-terminal residues was not observed, which suggests that the C-terminal in NAO is disordered. The model was refined to an overall R_{cryst} of 20.1 % and R_{free} of 23.2 % to 2.2 Å resolution (5% of the total reflections were not used in the refinements; (Brünger,

1992). The general model quality was judged form the estimated standard uncertainty values also known as diffraction precision index, calculated using *REFMAC5* (Cruickshank, 1999) (Table 4.3) and temperature factor plots (Figure 4.2). For detailed refinement statistics see (Table 4.4) The plots of the backbone torsion angles Φ, Ψ (Ramakrishanan & Ramachandran, 1965) for all the subunits of the homotetramer was calculated using *PROCHECK* analysis (Laskowski *et al.*, 1993). All the residues lie either in the most or additionally allowed region except for Thr 90 from each subunit (Figure 4.1).

Table 4.3 Diffraction precision index (DPI) calculated based on R_{cryst}, R_{free} and molecular likelihood value criteria.

REFMAC DPI (R_{cryst})[a]	0.22 Å
REFMAC DPI (R_{free})[b]	0.18 Å
REFMAC ML[c]	0.12 Å

[a] $DPI(R_{cryst})$ is given by $(N_i/p)^{1/2}C^{-1/3}\ R_{cryst}\ D_{min}$

[b] $DPI(R_{free})$ is represented by $(N_i/n_{obs})^{1/2}C^{-1/3}\ R_{free}\ D_{min}$

where N_i = number of atoms, corrected for scattering power; P = (number of observations – number of refined parameters); C = completeness of data at resolution D_{min}; D_{min} = resolution of data used in refinements; n_{obs} = number of reflections used for refinement, not including the reflections used in calculating R_{free}.

[c]DPI(ML) is based on the agreement of Fo and Fc by the resolution shells.

Table 4.4 Refinement statistics for ES*

Number protein residues	1720
Number water molecules	694
Resolution range (Å)	40 -2.2
Number of reflections	130649
R_{cryst} (%)[a]	20.16
R_{free} (%)	23.22
Ramachandran Statistics	
Most favored	91.8%
Additional allowed	7.7%
Generously allowed	0.2%
Disallowed	0.3%
Average B values (Å^2)	
Protein	33.95
Water	37.62
FAD	34.16
Ligand	52.65 (c atoms); 75.7 (nitro group) (100 % occupancy) (46.4 c atoms); 66.2(nitro group) (80 % occupancy)
R.m.s.d. from ideal geometry	
Bond lengths (Å)	0.006
Bond angles (^{o})	1.1

[a] R_{cryst} is calculated as $\{\Sigma_{h,k,l} \mid (|F_{obs}| - |F_{calc}|) \mid \} / (\Sigma_{h,k,l} |F_{obs}|)$; R_{free} was calculated using 5% of the total reflections collected at the three different wavelengths but not used for refinements.

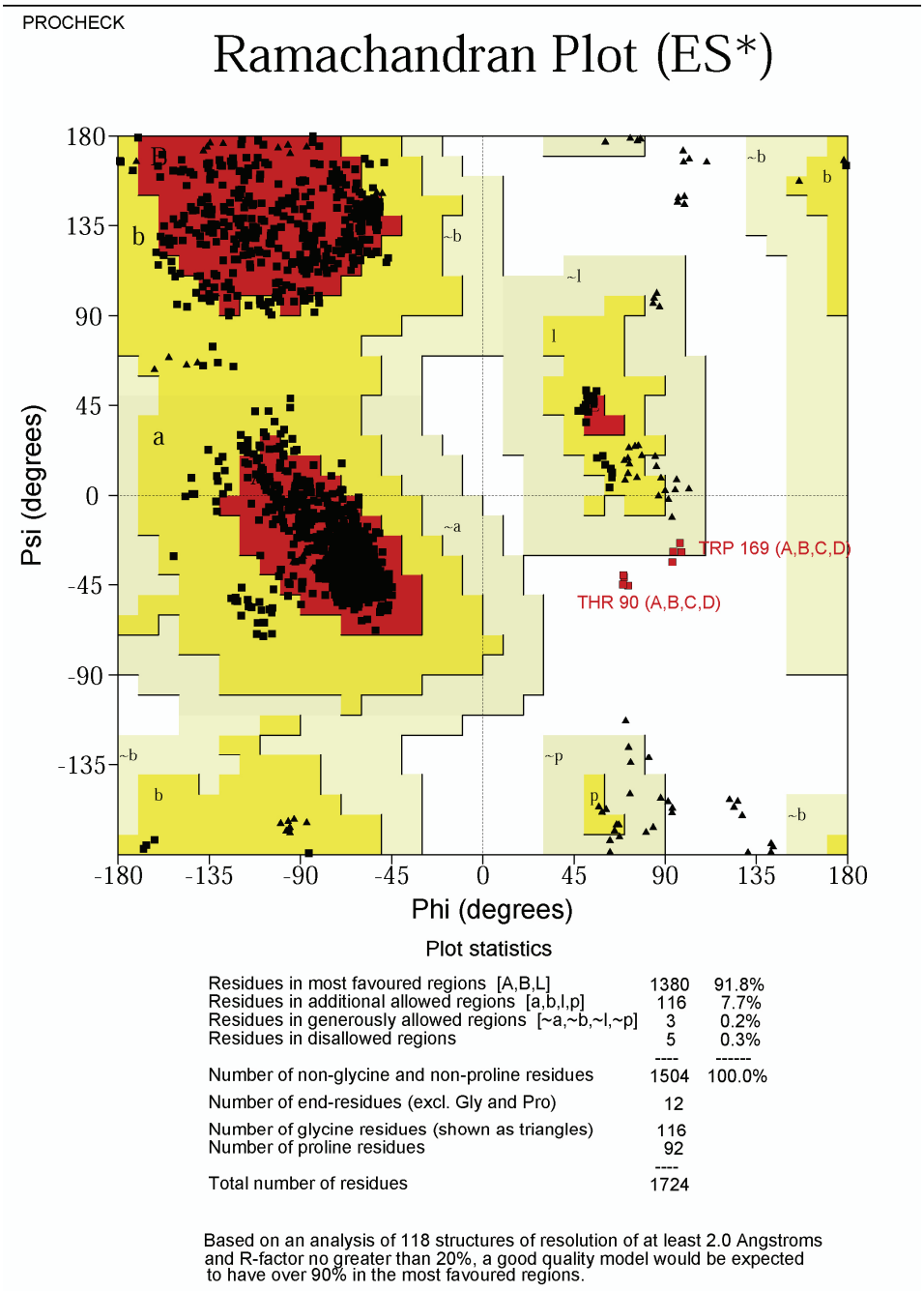


Figure 4.1 Ramachandran analysis for ES* homotetramer. The analysis was done using *PROCHECK* v.3.5.4 (Laskowski *et al.*, 1993).

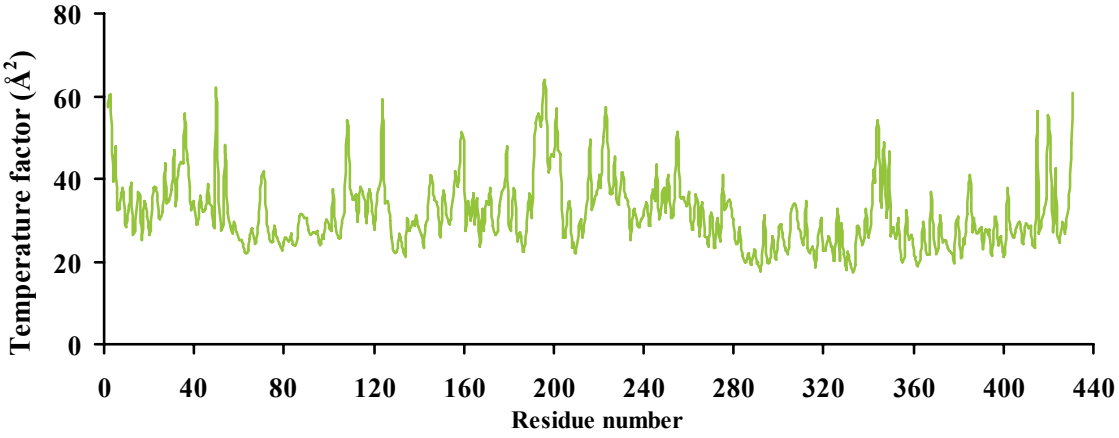


Figure 4.2 Variation in average temperature factor values along the polypeptide chain of ES*. The average B factor for each residue was calculated for all the chains in the ES*form of NAO. The calculations were done using *CNS* v.1.1.

4.4 Data Collection Statistics for Oxidized NAO (form I, II and III)

NAO has been crystallized in three crystal forms as described in the experimental section. In each case the crystals were bright yellow suggesting the presence of oxidized FAD in the active site. Each crystal form diffracts X-rays to at least 3.2 Å resolution and the data collection statistics are presented in (Table 4.5, 4.6). Crystal form I is remarkable because the trigonal space group has a 485 Å unit cell edge along *c* and typically display relatively low mosaic spread. However, the unit cell parameters pushed the limits of the data collection facilities available at several synchrotron X-ray facilities. Before a complete data set to 2.07 Å resolution was collected at BIOCARS (14-BMC) several attempts to collect a complete data set were unsuccessful (Table 4.6). Solvent content and Matthews coefficient analysis (Kantardjieff & Bernhard, 2003) suggested 1.5 NAO homotetramers in the asymmetric unit (Table 4.6). Slight alteration of the crystallization

conditions yielded crystals in different space group (form II). Analysis of X-ray diffraction data revealed that a different space group and unit cell parameters for the NAO crystals in form II. The solvent content and Matthews coefficient analysis (Matthews, 1968) suggested that there were two NAO homotetramers in the asymmetric unit (Table 4.5).

Crystal form III grew with SeMet enriched NAO and conditions that included DTT, but otherwise typically yielded crystal form 1. The resulting space group was trigonal and systematic absences suggested either $P3_1$ or $P3_2$, but the *c* unit cell edge was significantly shorter (Table 4.5). The solvent content and Matthews coefficient analysis (Matthews, 1968) suggested that there were two NAO homotetramers in the asymmetric unit (Table 4.5) and a higher solvent content. Despite the presence of arsenic in the sodium cacodylate buffer system, a three wavelength MAD was successfully collected. Although preliminary analysis of the MAD data revealed good phasing statistics, the MAD data was not used for final structure solution of E^{ox}. Significant computational cycles in the lab were used for locating 52 selenium sites in the MAD data set collected from SeMet-ES* crystals. Since ES* crystal structure was solved before 78 selenium sites could be located in the MAD data sets collected for form III, the three dimensional structure of oxidized NAO was solved by using molecular replacement methods.

Table 4.5 Data collection statistics for oxidized NAO (Form II and III)

Crystal Form	II	III (Se-Met)	III (Se-Met)	III (Se-Met)
X-ray Source	SER-CAT	SER-CAT	SER-CAT	SER-CAT
Beamline	22-ID	22-ID	22-ID	22-ID
Wavelength (Å)	1.0032	0.9792 (Peak)	0.9794 (Inflection)	0.9686 (Remote)
Detector	marccd 165	marccd 225	marccd 225	marccd 225
Resolution Range (Å) [a]	35 – 3.2 (3.4 – 3.2)	50 – 2.8 (2.9 – 2.8)	50 – 2.8 (2.9 – 2.8)	50 – 2.8 (2.9 – 2.8)
Mosaic Spread (°)	~ 0.6	~ 0.3	~ 0.3	~ 0.3
Space group	$P2_12_12_1$	$P3_1$ or $P3_2$	$P3_1$ or $P3_2$	$P3_1$ or $P3_2$
Unit Cell (Å), *a*	147.3	108.9	108.8	109.0
b	153.5	108.9	108.8	109.0
c	169.5	342.5	343.2	342.3
V_M (A^3/Da)	2.5	3.0	3.0	3.0
Solvent Content (%)	50	59	59	59
Enzyme (α) / Asym. Unit	8	8	8	8
Total Reflections	239001	767349	787263	690400
Unique Reflections	55534	111892	111132	111156
Multiplicity [a]	4.3 (2.7)	6.9 (5.2)	7.1 (6.4)	6.2 (4.5)
Completeness (%) [a]	92.6 (92.6)	99.9 (99.6)	99.9 (99.8)	99.9 (99.8)
R_{sym} [a,b]	0.073 (0.29)	0.126 (0.39)	0.126 (0.44)	0.147 (0.37)
$I/\sigma(I)$ [a,c]	7.8 (2)	11.4 (3.7)	11.8 (3.5)	11.5 (3.2)

[a] Values for the highest resolution shell of data are given in parentheses. [b] R_{sym} (I) gives the average agreement between the independently measured intensities such as $\Sigma_h\Sigma_i |I_i - I| / \Sigma_h\Sigma_i I$, where I is the mean intensity of the i observations of reflection h. [c] $I/\sigma(I)$ is the root-mean-square value of the intensity measurements divided by their estimated standard deviation.

Table 4.6 Data collection statistics for oxidized NAO (Form I)

Crystal Form	I	I	I	I
Timeline	Aug. 2002	Feb. 2003	March 2003	Aug. 2003
X-ray Source	SER-CAT	SER-CAT	NSLS	BIOCARS
Beamline	22 - ID	22 - ID	X26C	14-BMC
Wavelength (Å)	1.000	1.0332	1.0008	0.900
Detector	Mar165	Mar 165	Quantum 4R	Quantum 4R
Active area of Detector (mm)	162 X 162	162 X 162	188 X 188	188 X 188
Detector Distance	300	400	400	400
Resolution Range (Å) [a]	30 – 4.0 (3.85-4.0)	30 – 2.25 (2.15 – 2.25)	50 - 2.5 (2.6 - 2.5)	50 - 2.07 (2.15 -2.07)
Mosaic Spread (°)	0.38	0.32	0.4	0.36
Space group	$P3_221$	$P3_221$	$P3_221$	$P3_221$
Unit Cell (Å), *a*	103.72	103.14	103.4	103.82
b	103.72	103.14	103.4	103.82
c	487.45	484.09	485.1	485.4
V_M (A^3/Da)	2.6	2.6	2.6	2.6
Solvent Content (%)	53	53	53	53
Enzyme (α) / Asym. Unit	6	6	6	6
Total Reflections	81400	200379	265114	570374
Unique Reflections	38584	169489	199169	371128
Multiplicity [a]	2.1(2.0)	1.2 (1.0)	1.3 (1.0)	1.53 (1.0)
Completeness (%) [a]	78 (77.5)	65.1 (25.4)	89.1 (77.1)	93.5 (80.0)
R_{sym} (%) [a,b]	10.6 (14.0)	(30.4)	9.3 (18)	13.9 (30.1)
$I/\sigma(I)$ [a,c]	6.0 (4.8)	11.2 (4.2)	14.3 (8.3)	12.6 (3.0)

[a] Values for the highest resolution shell of data are given in parentheses. [b] R_{sym} (I) gives the average agreement between the independently measured intensities such as $\Sigma_h\Sigma_i |I_i - I| / \Sigma_h\Sigma_i I$, where I is the mean intensity of the i observations of reflection h. [c] $I/\sigma(I)$ is the root-mean-square value of the intensity measurements divided by their estimated standard deviation.

4.5 The Quality of Oxidized NAO Crystal Structure

The final model consists of 1.5 homotetramers per asymmetric unit. A total of 2580 residues were built in the asymmetric unit, 430 per subunit. The electron density for eight C-terminal residues was not observed, which suggests that the C-terminal in NAO is disordered. The model was refined to an overall R_{cryst} of 18.8 % and R_{free} of 22.5 % to 2.07 Å resolution (5% of the total reflections were not used in the refinements; (Brünger, 1992). The general model quality was judged form the estimated standard uncertainty values also known as diffraction precision index, calculated using *REFMAC5* (Cruickshank, 1999) (Table 4.7) and the temperature factor plot (Figure 4.4). For detailed refinement statistics see (Table 4.8). The plots of the backbone torsion angles Φ, Ψ (Ramakrishanan & Ramachandran, 1965) for all the subunits of the homotetramer were calculated using *PROCHECK* analysis (Laskowski *et al.*, 1993). All the residues lie either in the most or additionally allowed region except for Thr 90 from each subunit (Figure 4.3).

Table 4.7 Diffraction precision index (DPI) calculated based on R_{cryst}, R_{free} and molecular likelihood value criteria.

REFMAC DPI (R_{cryst})[a]	0.20 Å
REFMAC DPI (R_{free})[b]	0.17 Å
REFMAC ML[c]	0.11 Å

[a] $DPI(R_{cryst})$ is given by $(N_i/p)^{1/2}C^{-1/3}\ R_{cryst}\ D_{min}$

[b] $DPI(R_{free})$ is represented by $(N_i/n_{obs})^{1/2}C^{-1/3}\ R_{free}\ D_{min}$

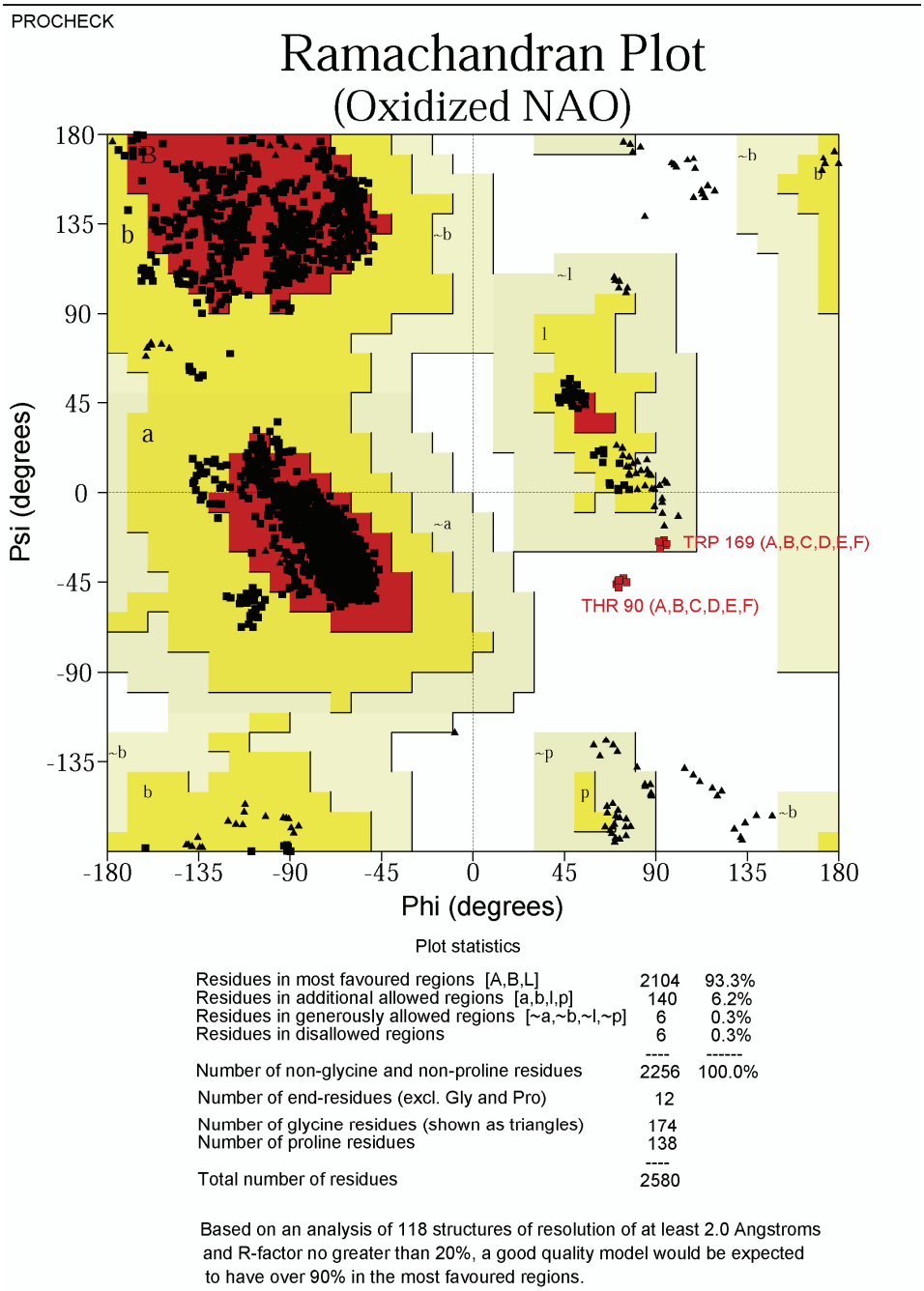
where N_i = number of atoms, corrected for scattering power; P = (number of observations – number of refined parameters); C = completeness of data at resolution D_{min}; D_{min} = resolution of data used in refinements; n_{obs} = number of reflections used for refinement, not including the reflections used in calculating R_{free}.

[c] DPI(ML) is based on the agreement of Fo and Fc by the resolution shells.

Table 4.8 Refinement statistics for oxidized form of NAO

Number protein residues	2580
Number water molecules	1136
Resolution range (Å)	50 – 2.07
Number of reflections	171484
R_{cryst} (%)[a]	18.76
R_{free} (%)	22.51
Ramachandran Statistics	
Most favored	93.3%
Additional allowed	6.2%
Generously allowed	0.3%
Disallowed	0.3%
Average B values (Å^2)	
Protein	26.73
Water	31.71
FAD	19.81
Ligand (Spermine) - 4 molecules	46.31
Ligand (Glycerol) - 14 molecules	50.23
Ligand (PEG 4000) - 2 molecules	51.45
R.m.s.d. from ideal geometry	
Bond lengths (Å)	0.013
Bond angles (^{o})	1.4

[a] R_{cryst} is calculated as $\{\Sigma_{h,k,l} \mid (|F_{obs}| - |F_{calc}|) \mid \} / (\Sigma_{h,k,l} |F_{obs}|)$; R_{free} was calculated using 5% of the total reflections collected but not used for refinements.



Plot statistics

Residues in most favoured regions [A,B,L]	2104	93.3%
Residues in additional allowed regions [a,b,l,p]	140	6.2%
Residues in generously allowed regions [~a,~b,~l,~p]	6	0.3%
Residues in disallowed regions	6	0.3%
Number of non-glycine and non-proline residues	2256	100.0%
Number of end-residues (excl. Gly and Pro)	12	
Number of glycine residues (shown as triangles)	174	
Number of proline residues	138	
Total number of residues	2580	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

Figure 4.3 Ramachandran analysis for oxidized form of NAO. The analysis was done using *PROCHECK* v.3.5.4 (Laskowski *et al.*, 1993).

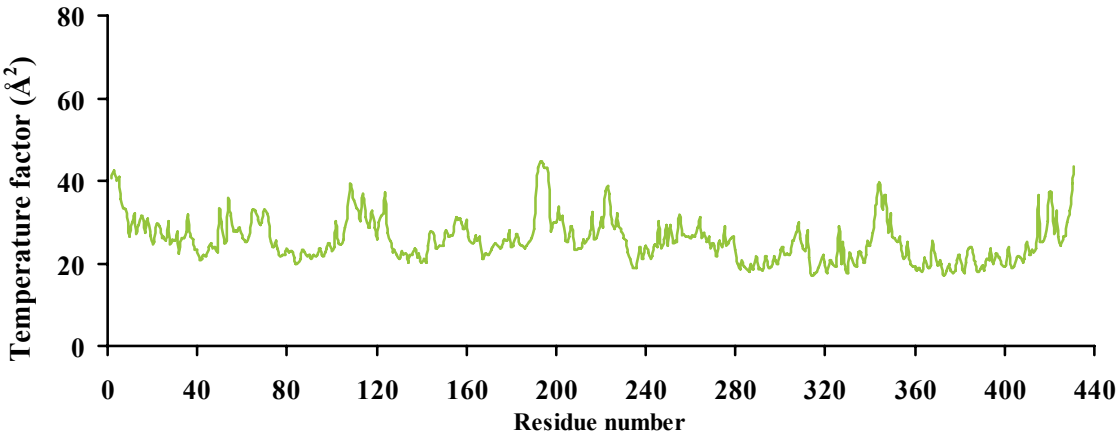


Figure 4.4 Variation in the average temperature factor values along the polypeptide chain. The average B factor for each residue was calculated for all the chains in the oxidized form of NAO. The calculations were done using *CNS* v.1.1.

4.6 Structural Comparison of Different Reaction States of NAO

The plot of average solvent accessible surface area along the polypeptide chain in all the three different forms of NAO is very similar (Figure 4.5). The continuous regions of low solvent accessibility occur between 73-105, 277-293, 327-340, 352-380 and 394-414. Most of the residues in these regions correspond to the N5 access substrate access channel and the active site loop. The highest solvent accessible surface area (17 – 20 $Å^2$) is observed for residues numbers 36, 64, 157, 222, 344, 420 and 431.

The structural overlap between ES* and oxidized NAO homotetramer gave a low r.m.s.d. of 0.3 Å. The r.m.s deviation plots along the polypeptide chain between the different reaction states of NAO show small shifts in the protein backbone (Figure 4.6 *a,*

b, c). The large shifts in the side chain conformations are correlated with the high solvent accessible surface areas, which is expected since solvent exposed side chain can adopt different conformations. The r.m.s deviations are more meaningful for the residues that have low solvent accessibility. The residues that are solvent exposed are more likely to have different side chain conformations in comparison to the residues that are not involved in crystal packing and are buried inside the protein. For these reasons, a 1.26 Å shift in Asp402 side chain can be stated with greater confidence in comparison to a 4 Å shift observed for the side chains of Lys36, Glu 344 (Figure 4.6 *a, b, c*).

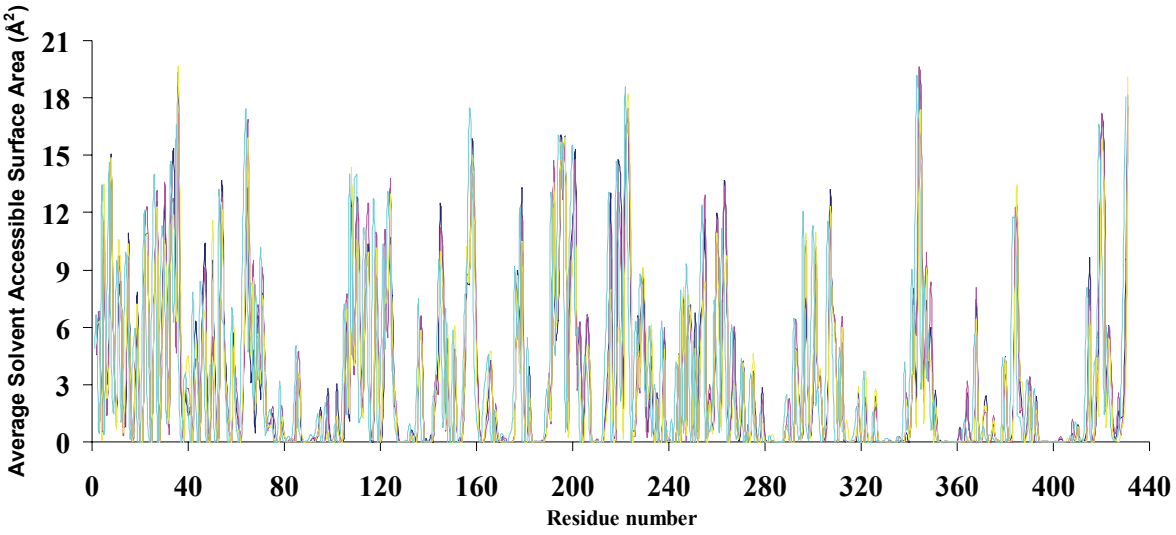


Figure 4.5 Average solvent accessible surface area of the residues in different reaction states of NAO. The uninhibited state of NAO, E^{ox} (chain A) is shown in green; The inhibited state of NAO, EI (chain B) is shown in magenta; The N5-FAD adduct of NAO, ES* (chain A) as shown in yellow.

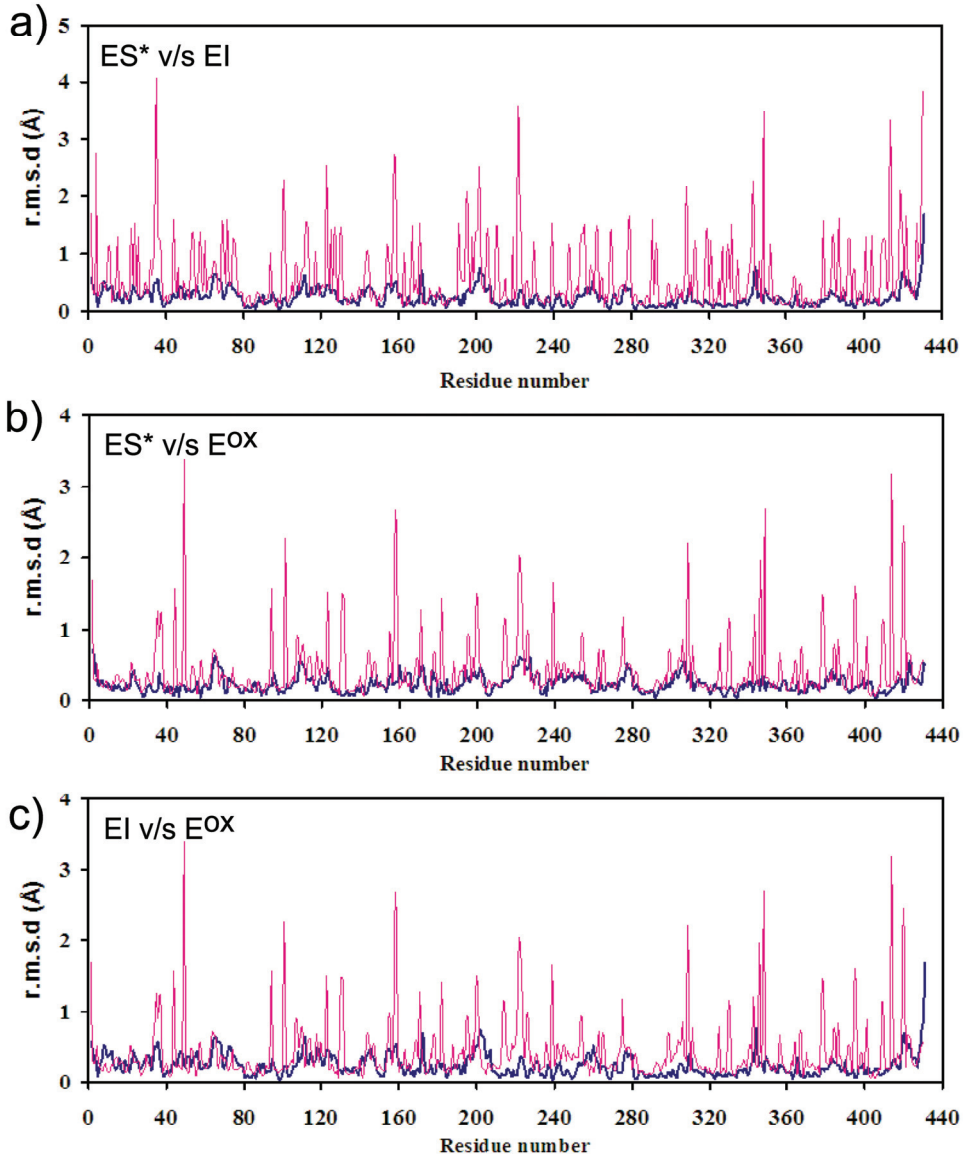


Figure 4.6 a) The mean positional variation of the side chain (magenta) and the main chain (violet) residues of ES* (chain B) versus the residues in the inhibited state of NAO, EI (chain B); b) The mean positional variation of the side chain (magenta) and the main chain (violet) residues of ES* (chain A) versus the residues in the uninhibited state of NAO, E^{ox} (chain A); c) The mean positional variation of the side chain (magenta) and the main chain (violet) residues of EI (chain B) versus the residues in the uninhibited state of NAO, E^{ox} (chain A). Individual chains were superimposed using *SUPERPOSE* routine in *CCP4* (Krissinel & Henrick, 2004).

5 CRYSTAL STRUCTURES OF NITROALKANE OXIDASE

[††]Nitroalkane oxidase (NAO) from *Fusarium oxysporum* catalyzes the oxidation of the neutral nitroalkanes to the corresponding aldehydes or ketones with production of H_2O_2 and nitrite. The flavoenzyme is a new member of the acyl-CoA dehydrogenase (ACAD) family, but it does not react with acyl-CoA substrates. We present the 2.2 Å resolution crystal structure of NAO trapped during turnover of nitroethane as a covalent N5-FAD adduct (ES*). The homotetrameric structure of ES* was solved by MAD phasing with 52 Se-Met sites in an orthorhombic space group. The electron density for the N5-(2-nitrobutly)-1,5-dihydro-FAD covalent intermediate is clearly resolved. The structure of ES* was used to solve the crystal structure of oxidized NAO (E^{ox}) at 2.07 Å resolution. The *c* axis of trigonal space group for E^{ox} is 485 Å and there are 1½ holoenzymes in the asymmetric unit. Four of the active sites contain spermine (EI), a weak competitive inhibitor and two are empty (E^{ox}). The structures of E^{ox}, EI and ES* reveal a hydrophobic channel that extends from the exterior of the protein to the *re*-face of the FAD which terminates at the N5-FAD position. The structures show that Asp402 is located at the end of the substrate access channel. Moreover, the residue is in the correct position to serve as the active site base, where it is proposed to abstract the α-proton from neutral nitroalkane substrates. The protein fold for NAO and various members of the ACAD family overlay with root mean square deviations between 1.7 to 3.1 Å. The homologous region typically spans more than 325 residues and includes Glu376, which is

---

[††] This chapter has been reformatted from the NAO manuscript prepared for Nat. Str. Biol. (Nagpal *et al.*, 2005)

the active site base in the prototypical member of the ACAD family. However, NAO and the ACADs exhibit differences in hydrogen the bonding between the respective active site base, substrate molecules and the FAD. These likely differentiate NAO from the homologs and consequently, are proposed to result in the unique reaction mechanism of NAO.

5.1 Overview of the NAO Structures

The trapped intermediate of NAO, ES*, crystallized at 4 °C in space group $P2_12_12_1$ with one homotetramer per asymmetric unit (V_M 3.4; 63% solvent). The ES* structure has 222 local symmetry relating the four subunits to each other (Figure 5.1). The tetrahedron-shaped holoenzyme is approximately 100 Å wide (measured between C_α of Gly158a and Gly158b[‡‡]; and approximately 80 Å tall (measured between C_α of Thr34b and Thr34c). The oxidized form of NAO, E^{ox}, crystallized in space group $P3_221$ from slightly different conditions that include spermine tetrahydrochloride (Nagpal *et al.*, 2004). The volume of the trigonal unit cell is greater than 4,570,000 $Å^3$ and contains six monomers per asymmetric unit (V_M 2.6; 53% solvent). These are arranged as one complete homotetramer and one-half of another holoenzyme. However, the additional dimer is oriented with a local two-fold symmetry axis coincident with a crystallographic two-fold axis.

[‡‡] Residue naming convention – residue number followed by chain ID

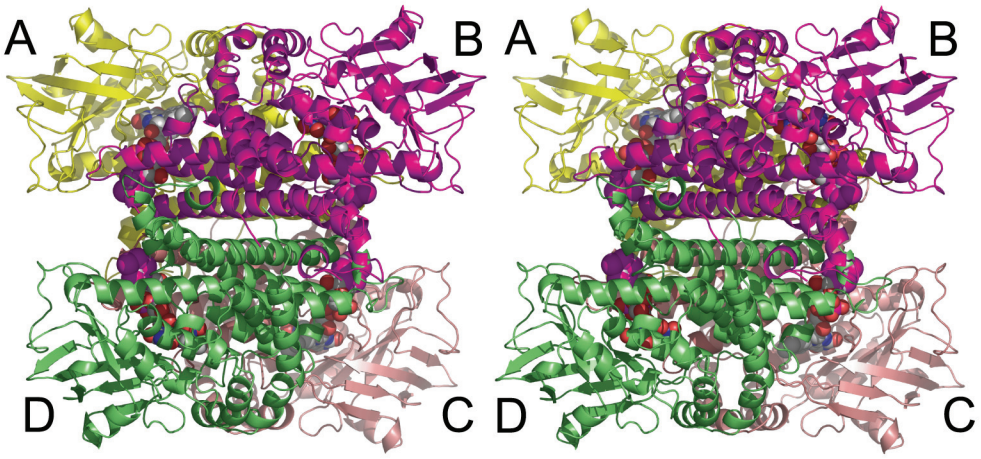


Figure 5.1 A divergent stereo view of the NAO homotetramer in the ES structure. Each subunit is depicted as a differently colored C_α ribbon trace and the FAD covalent adducts are shown in CPK atoms within each subunit. An A:B dimer can be considered as consisting of a purple and yellow pair of subunits.

Biochemical studies (Gadda & Fitzpatrick, 1998) and our crystal structures suggest that NAO is a homotetramer comprised of a dimer of dimers. Subunit A forms extensive interfaces with subunits B and C. In contrast, only the side chains of Gln314a, Ser315a and Asp318a form hydrogen bonds across a two-fold axis with residues from subunit D. The buried surface area at the A:C dimer interface is slightly greater (~ 5270 $Å^2$; 16.5% of the total solvent accessible surface area for a subunit) than that between the A:B dimer (~ 4560 $Å^2$; 14% of the total). The A:C dimer and B:D dimer are stabilized by several types of interactions. For example, residues from subunit A along helix K (Figure 5.2, 5.3) and the C-terminal region form hydrogen bonds with residues from helix H and G of the subunit C. In addition, four salt bridges also stabilize the A:C interface: Lys324a-Glu353c, Arg410a-Asp318c, Arg410a-Asp322c and Arg411a-Asp318c. The active site within subunit A contains residues from chain A as well as few of the residues

from subunit B (discussed further below). Therefore, the homotetramer can be described as an A:B dimer with a C:D dimer interface.

The secondary and tertiary structures of each NAO subunit is comprised of two α-helical domains (approximately residues 1-125 and 260-438) and a central β-sheet domain (approximately residues 125-260) (Figure 5.2, 5.3). The α-helical domains form the core of the holoenzyme and the β-sheet domains are located on the exterior, approximately at the points of the tetrahedron shape. Each subunit in the two asymmetric units of the oxidized NAO and the ES* structures have very similar structures (r.m.s.d. of all C_α atoms ~ 0.25 Å), despite the fact that non-crystallographic symmetry was not restrained during the refinements.

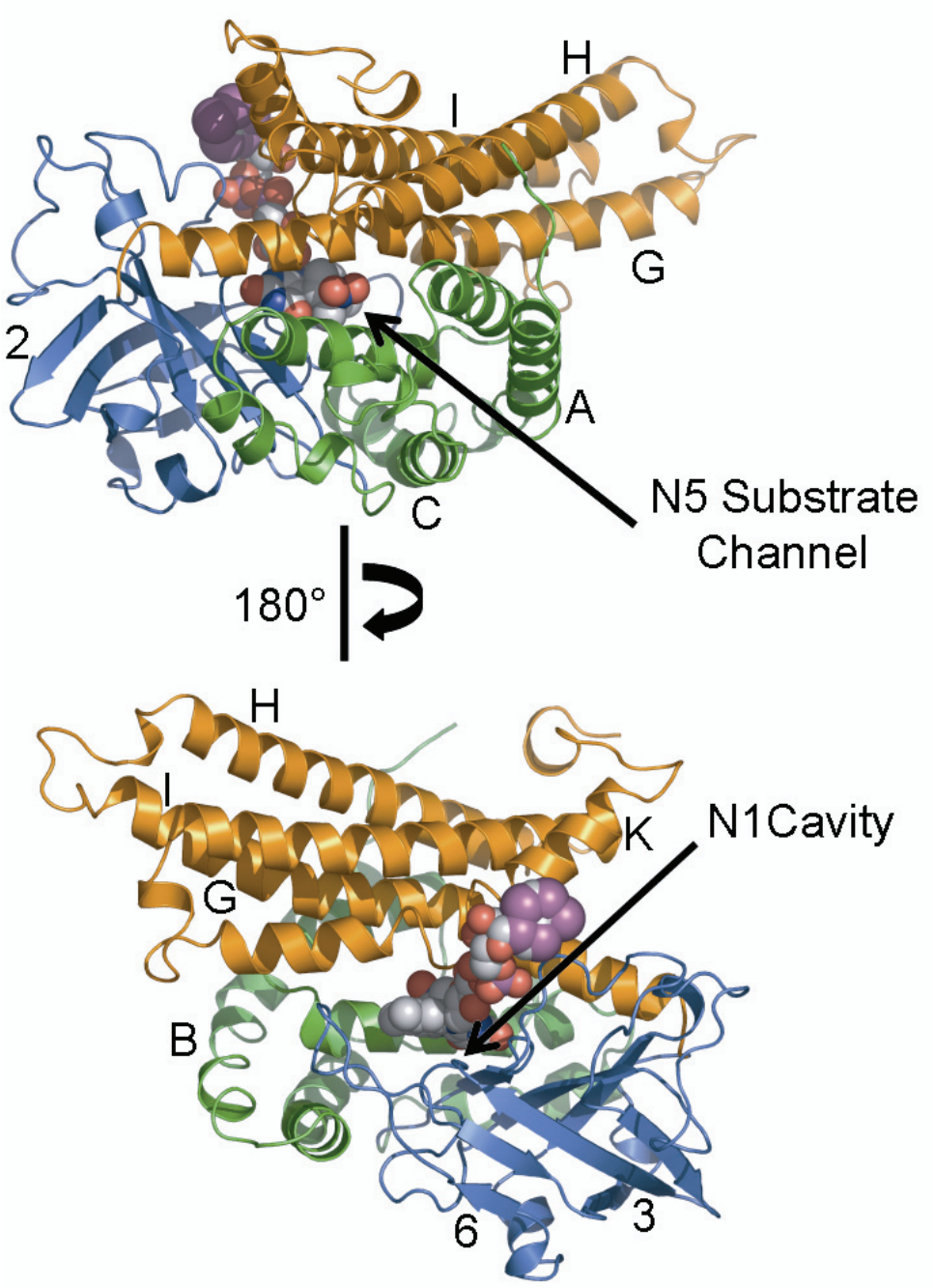


Figure 5.2 A ribbon drawing of one NAO subunit showing the β sheet domain (blue) sandwiched between the N-terminal α domain (green) and the C-terminal α domain (gold). The FAD is shown as CPK atoms with the N5 substrate access channel and N1 cavity indicated by the arrows

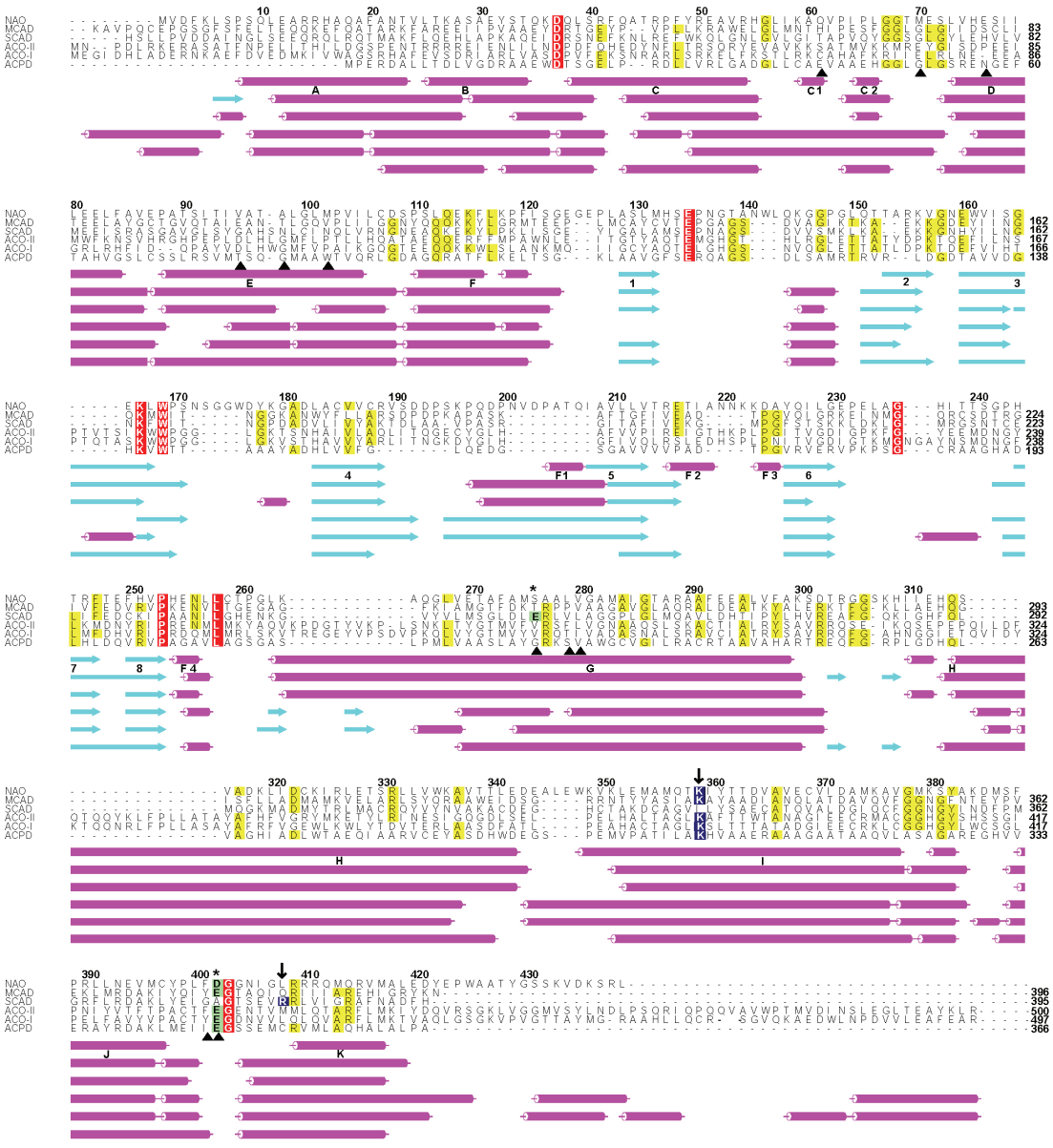


Figure 5.3 Amino acid sequence and 2° structural and alignments of NAO and several homologs. The residue numbering for NAO is on the top and for the homologs on the right hand side. Naming of the 2° structural elements are for the NAO structure. The 2° structural assignments are shown in the same order as the 1° sequences. MCAD, medium chain acyl-CoA dehydrogenase (3mde); SCAD, short chain isovaleryl acyl-CoA dehydrogenase (1ivh); ACO-II, acyl-CoA oxidase II from rat liver peroxisomes (1is2); ACO-I, acyl-CoA oxidase I from *Arabidopsis thaliana* (1w07); Acyl-ACP, acyl-ACP dehydrogenase from the FK520 polyketide biosynthetic pathway (1r2j). The active site residues lining the N5 substrate access channel in NAO are indicated with filled triangles. Residues that are 100% conserved are highlighted in red and those present in at least 4/6 are in yellow. The active site bases are all donated from the J-K loop, except where noted (shown in green). The conserved lysine that stabilizes the J-K loop is from α-helix I, except where noted (shown in blue). The ACO-II and ACO-I amino acid sequences are truncated to 500 and 497 residues respectively. The alignment was prepared by using *CLUSTALW* and *ALSCRIPT* (Higgins *et al.*, 1994; Barton, 1993).

5.2 Architecture of the Active Site

Each NAO subunit binds one FAD non-covalently in an extended conformation (Figure 5.2, 5.4). The isoalloxazine ring is located between the β sheet domain on the *si*-face and on the *re*-face by α-helices E, G, F1 and the JK loop. The FAD in subunit A also packs against subunit B in several regions. For example, the turn between Ala376b to Met379b, the only counter ion to the FAD phosphate moieties is Arg304b, and His313b π-stacks with the adenine ring of the FAD. The adenine moiety of the FAD extends to the dimer-dimer interface, where it hydrogen bonds with the side chains of Gln314 of the subunit C and Asn141 from subunit A.

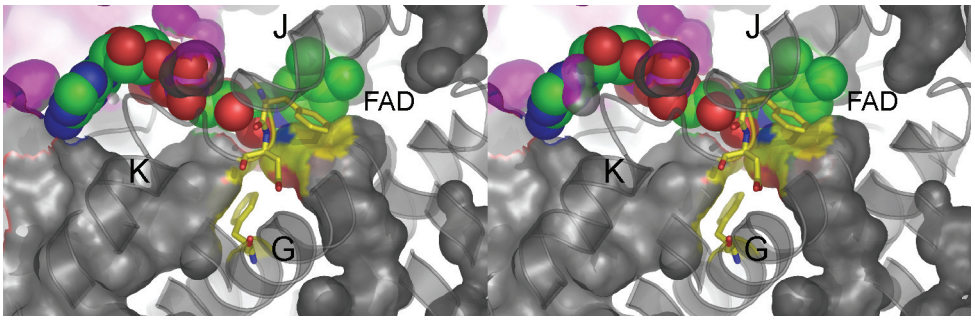


Figure 5.4 A divergent stereo view of the solvent accessible surfaces around the FAD in subunit A of the E^{ox} active site. The FAD is shown as CPK atoms with carbon, nitrogen, oxygen and phosphorus atoms colored green, blue, red and magenta, respectively. The active site base, Asp402, as well as Phe401 and Phe273 are shown as sticks with carbon atoms colored yellow. The surfaces are colored according to the atoms that form the boarder, with gray surface in subunit A and purple surfaces from subunit B. The N5 channel extends from the bottom-right, to the right of α helix G and terminates under the Phe401-Asp402 residues. The N1 channel extends from the bottom-left, under the α helix K and terminates near Phe273 from α helix G.

The crystal structure of oxidized NAO contains six subunits in the asymmetric unit. The electron density maps indicate that four of the active sites contain spermine (EI), a weak competitive inhibitor (K_i ~200 mM at pH 8) and two are empty (E^{ox}) (Figure 5.5 a,b). The electron density for the covalent adduct in the ES* structure is clearly resolved and suggests full occupancy (Figure 5.5 c). Our crystal structures correlate well with previous biochemical studies and confirm the presence of Cys397, Tyr398 and Asp402 (Gadda *et al.*, 2000b; Gadda *et al.*, 2000a; Valley & Fitzpatrick, 2003b) in the active site. The active site Tyr398 was identified by tetranitromethane modification (Gadda *et al.*, 2000) and our structures show that it is located on the *re*-face of the FAD (Figure 5.6). Cys397 was also identified by chemical modification experiments (Gadda *et al.*, 2000) and also forms a portion of the active site cavity.

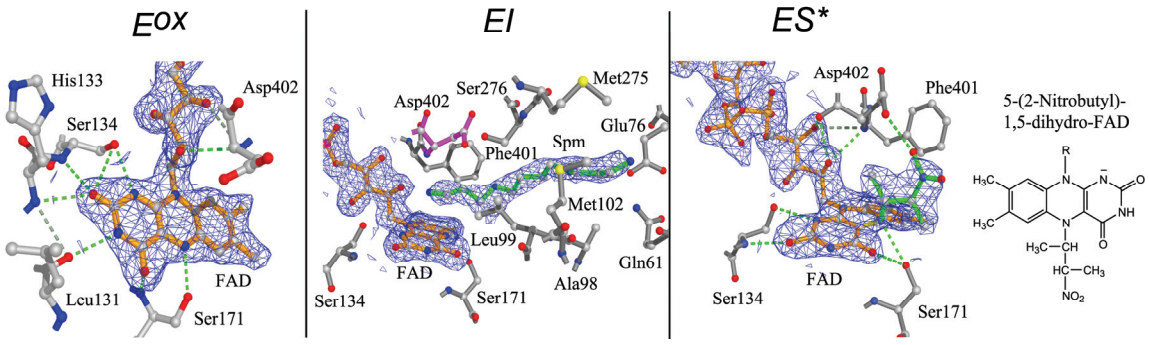


Figure 5.5 Examples of the electron density at the active site in the two structures of NAO. a) empty active site of oxidized NAO (E^{ox}) b) active site containing spermine inhibitor (EI). c) The covalent adduct trapped during turnover of nitroethane (ES*) and the corresponding chemical formula (Gadda *et al.*, 1997) are also shown. The $2mF_o - DF_c$ electron density maps are contoured at 1σ and calculated to 2.07 resolution for the E^{ox} and EI active sites and at 2.2 Å resolution for ES*.

Each of the structures reveals several solvent accessible or excluded cavities associated with the FAD (Figure 5.4). However, only about 16 $Å^2$ (4.5%) of the isoalloxazine ring surface area in E^{ox} is accessible to solvent (Figure 5.4). The largest solvent accessible channel extends from the exterior of the protein and terminates at the N1 atom, the ribityl portion of the FAD, and residue Phe273. We define this as the N1 channel. It is lined with several charged and hydrophilic residues, and contains several solvent molecules. The second channel also extends from the protein exterior and terminates at the FAD N5 atom and residues Phe401 and Asp402. We define this as the N5 channel and suggest that aliphatic substrates may enter the NAO active site through this channel. The 26 Å long channel is funnel-shaped and passes through the interface formed by the N-terminal and C-terminal α-helical domains. In the middle, an approximately 9 Å diameter constriction is formed by residues Val95 and Val280. The interior walls of the channel are bordered exclusively with the hydrophobic residues from α-helices C1, D, E, G and J-K loop. The total solvent accessible volume of the N5 channel is approximately 345 $Å^3$. The spermine molecules in the EI active sites occupy approximately 72% of that volume and exclude the FAD C4a and N5 atoms from solvent. This correlates with the preference for the hydrophobic nitroalkane substrates with each methylene group contributing ~2.6 kcal mol^{-1} of binding energy (Gadda & Fitzpatrick, 1999). The *re*-face of the isoalloxazine ring in the covalent adduct of ES* is similarly excluded from solvent, but one water molecule is hydrogen bonded to the O4 atom of the FAD in each structure. The exterior opening of the N5 channel is approximately 15 Å wide and is surrounded by the hydrophobic residues (Ile64, Pro65, Leu66, Ile105, Leu106, and Trp348) and charged residues (Glu71, Ser72, Thr271, Glu342, Glu344, and

Lys351). The presence of three negatively charged residues on the periphery of the substrate channel suggests, in part, a structural basis for NAO substrate discrimination against nitronate substrates. A third, solvent-excluded cavity is located parallel to the FAD *si*-face. It contains five ordered solvent molecules, but only the dimethylbenzene portion of the isoalloxazine ring boarders the dumbbell-shaped cavity. The rest of the cavity boundary is comprised of the side chains of several residues as well as portions of the protein backbone (Phe43, Ile92, Thr93, Trp169, Ser171, Asn172, Glu231, Pro232, Leu234, Thr240 and Gly242). The cavity is very close to the A:B subunit interface and only excluded from bulk solvent by the side chains of Leu168, Trp169 and Pro232.

The structures of E^{ox}, EI and ES* clearly show the location of Asp402, the essential active site base identified by several types of experiments (Gadda *et al.*, 2000b; Gadda *et al.*, 2000a; Valley & Fitzpatrick, 2003b). It is poised directly above the *re*-face of the FAD and the side chain also interacts with the guanidinio group of Arg409 via electrostatics and hydrogen bonds. Asp402 is located within a loop between the J and K α-helices (residues Pro399 to Asn405). As a result, the carbonyl oxygen atoms of Tyr398, Phe401, Gly403 and side chain oxygen of Asn405 are forced into close proximity to each other. The resulting negative electrostatic field is further enhanced by the C-terminal end of the J α-helix. These influences are compensated for by the positively charged N^{ε} atoms of Lys359 via hydrogen bonding and electrostatic interactions (Figure 5.6). Consequently, the J-K loop is properly orientated such that Asp402 is directly above the N5 atom of the FAD.

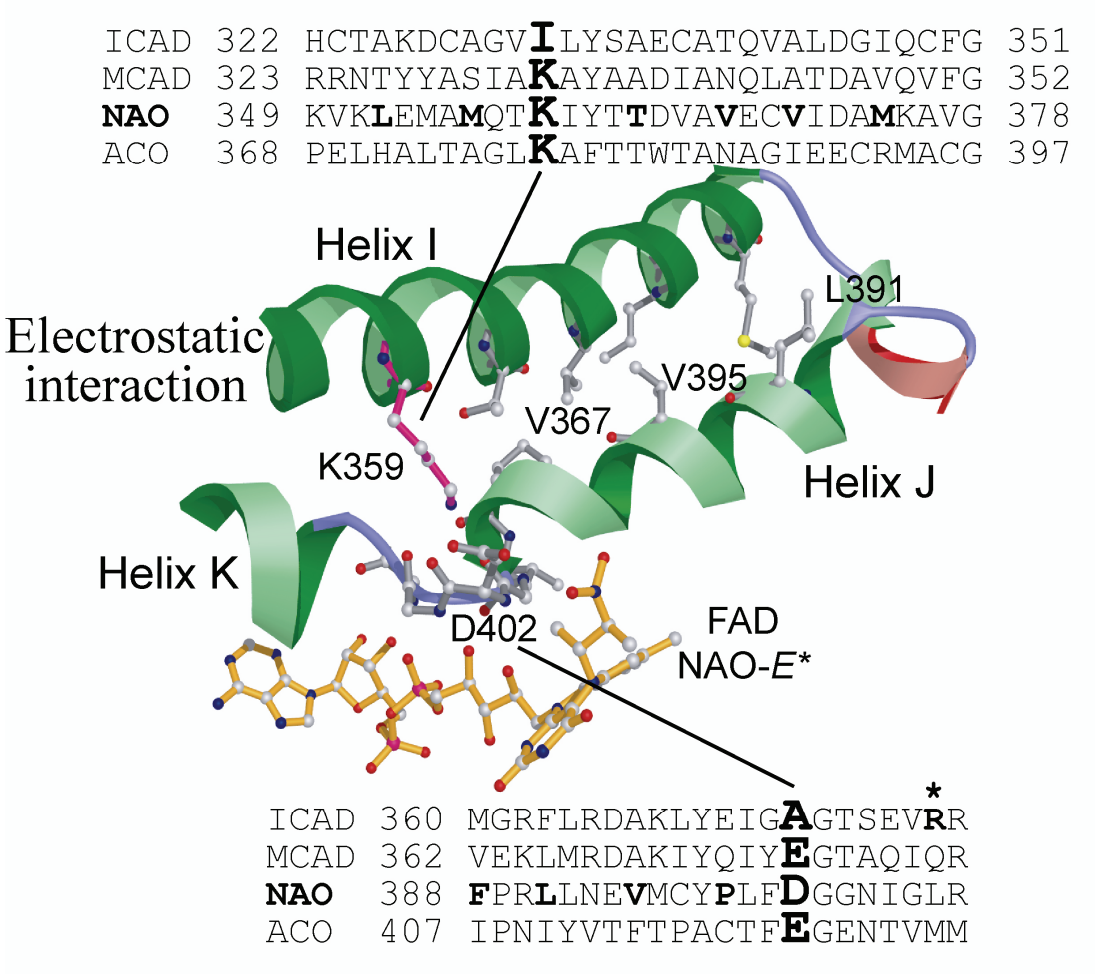


Figure 5.6 A view showing active site residues on JK loop. Lysine 359 forms electrostatic interactions with the backbone and side chain oxygens of the active site loop residues. Lysine is conserved throughout the ACAD family except for iso-valeryl acyl CoA dehydrogenase (ICAD) in which arginine 387 plays the same role.

The position of the FAD isoalloxazine ring is stabilized through several hydrogen bonds with the pyrimidine ring (Figure 5.7). This includes backbone atoms from Lue131, His133 and Ser134. In addition, the hydroxyl groups of Ser134 and Ser171 hydrogen bond with the FAD N1 and N5 atoms, respectively. Hydrophobic interactions with the dimethylbenzene ring are provided by Ile92, Trp169, Leu234, Phe401, and Met379b.

Moreover, there is a nearly perpendicular relationship between the aromatic ring of Phe401 and the FAD dimethylbenzene ring (Figure 5.4, 5.5). Such a T-shaped configuration is known to stabilize benzene and substituted benzene dimers by interactions between the electronegative π-orbitals and the electropositive edge of the aromatic ring (Sinnokrot & Sherrill, 2004; Pellett *et al.*, 2001) The electron density for the isoalloxazine rings in E^{ox} and EI indicate that they are distorted from planarity by approximately 15°, presumably by the combined influences of these interactions (Figure 5.8, Table 5.1). The piperizine ring in the ES* structure is also puckered at the N5 atom and consequently, yields a larger distortion (~21°). These features are consistent with a reduced FAD in the N5 adduct. Similar distortions of the isoalloxazine ring systems have also been observed recently for N5 and C4a covalent adducts of mono-amine and polyamine oxidases (Binda *et al.*, 2003; Binda *et al.*, 2001), which have structures that are otherwise completely unrelated to NAO. In each case, the deviation from planarity is principally along the N5-N10 axis, which bisects the essentially planar dimethylbenzene and pyrimidine ring systems.

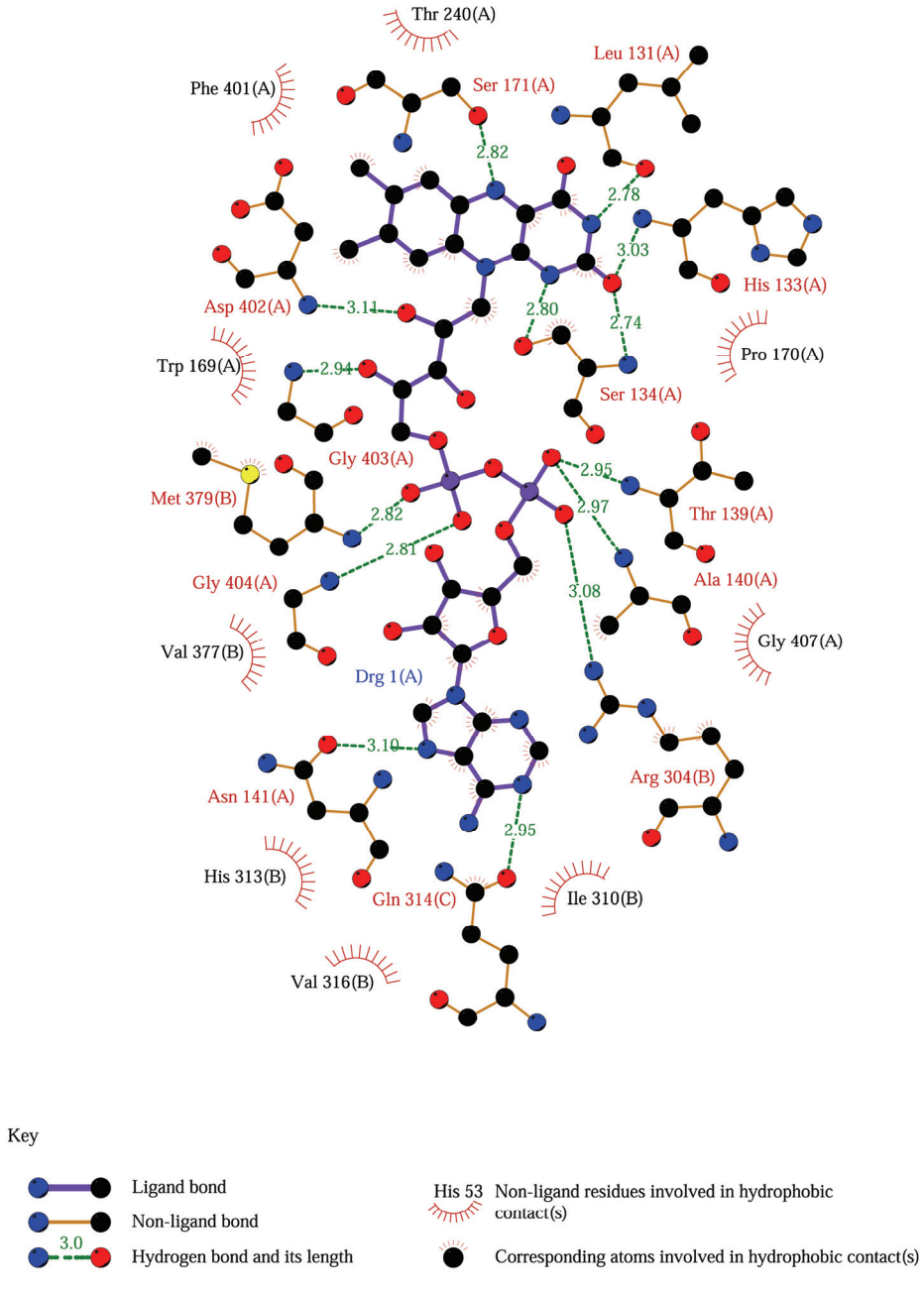


Figure 5.7 Schematic representation of the residues interacting with FAD in the active site of E^{ox}. The figure was prepared with *LIGPLOT* (Wallace *et al.*, 1995)

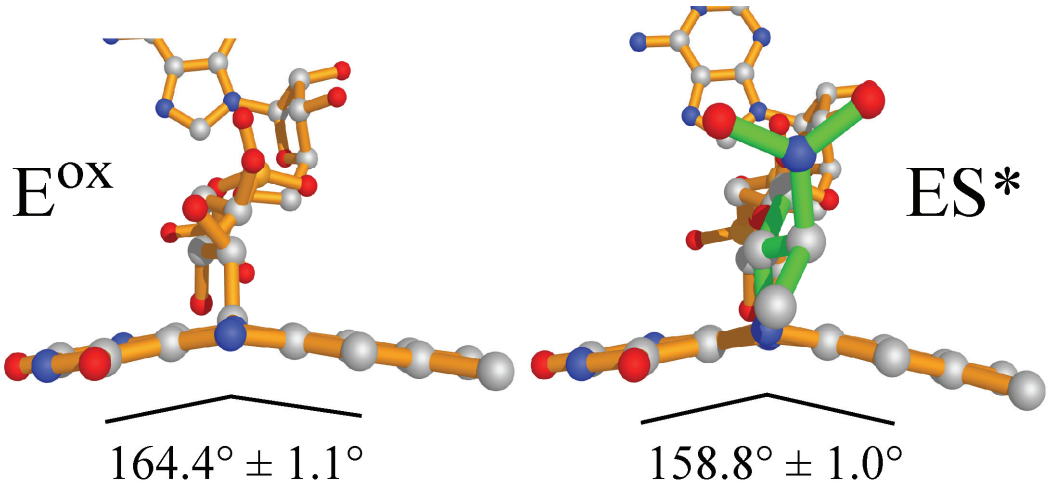


Figure 5.8 The isoalloxazine rings display deviations from planarity in both E^{ox} and ES* structures. The angle shown is the average of all the active sites in each asymmetric unit and depicts the C4-N5-C6 angle.

Table 5.1 Angles between the dimethylbenzene and pyrimidine rings of the FADs in the NAO structures [a]

	E^{ox}	ES*
C4-N5-C6 (°)	164.4 ± 1.1	158.8 ± 1.0
N1-N10-C9 (°)	164.9 ± 1.5	166.9 ± 0.7
Average (°)	164.6 ± 1.3	162.5 ± 4.8

[a] The values shown are for the average of the six or four independently refined active sites in each structure, respectively.

5.3 Comparisons of the E^{ox}, EI and ES* Active Sites

The overall polypeptide fold is very similar in the two structures of NAO (r.m.s.d. for C_α atoms of approximately 0.27 Å; 0.34 Å for 6876 backbone atoms in 429 residues per subunit). The structural overlay of homotetramers of oxidized NAO and the ES* complex reveals small shifts of the residues in the J-K loop and in several residues bordering the *si*-face. However, perhaps the functionally most important of these shifts are illustrated in (Figure 5.9, 5.10). For example, in the resting state of the enzyme (subunit A, E^{ox}), Asp402 forms many hydrogen bonds with several other residues and a solvent atom (Wat297). Most of these interactions are maintained in the EI and ES* complexes, but the three solvent molecules are displaced. Moreover, as illustrated, in subunit A of ES*, the orientation of Ser276 and Asp402 are both altered compared to the E^{ox} active site (Figure 5.9). Indeed, the side chain of Asp402 in subunit A of ES* has rotated almost 90° to form hydrogen bonds with the nitro group of the adduct and with the guanidinio group of Arg409. In addition, Ser276 has also rotated into an alternative orientation (Figure 5.11). In contrast, a comparison of subunit B from oxidized NAO and the ES* complex illustrates yet another set of conformations (Figure 5.10, 5.12). In this EI subunit, the spermine molecule has displaced the three solvent molecules. The protons from the terminal cationic nitrogen likely form hydrogen bonds with the FAD 2' OH ribityl chain and thus stabilize the binding of the inhibitor directly above the FAD *re*-face. However, the side chains of both Asp402 and Ser276 are in similar orientations within these subunits. Therefore, Asp402 has at least two possible orientations, which are coupled with the orientation of Ser276.

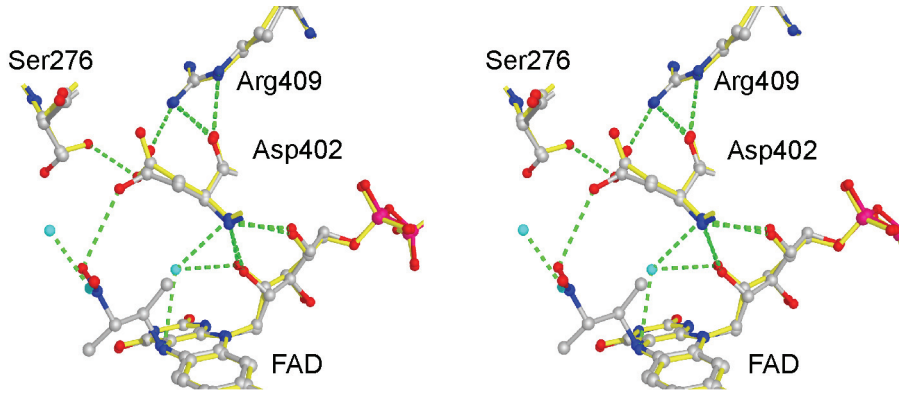


Figure 5.9 A divergent stereo view comparing E^{ox} (subunit A) with ES* (subunit A). Bonds for the E^{ox} structure are in yellow, whereas those for the ES* are in grey; hydrogen bonds are shown with dashed green lines. Three solvent molecules (cyan) in the E^{ox} structure are not present in the ES* complex. Note the alternative orientations and hydrogen bonding patterns associated with Asp402 and Ser276 for the two structures.

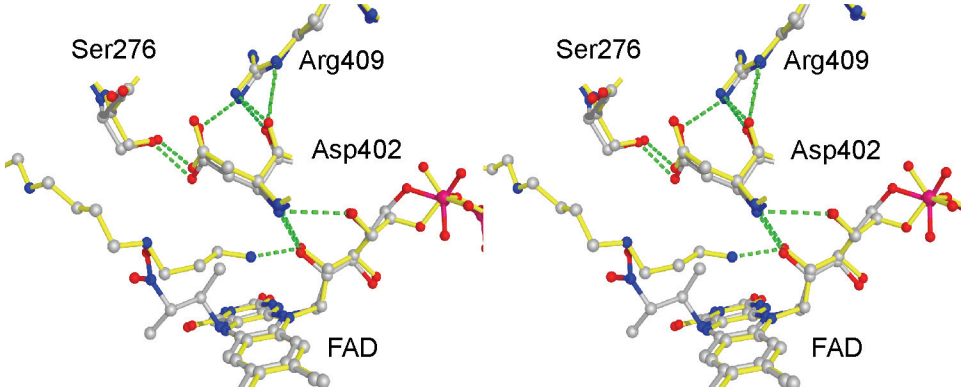


Figure 5.10 A divergent stereo view comparing EI (subunit B) with ES* (subunit B). The bonds for the EI complex are shown in yellow. Note that Asp402 and Ser276 have very similar orientations in each structure.

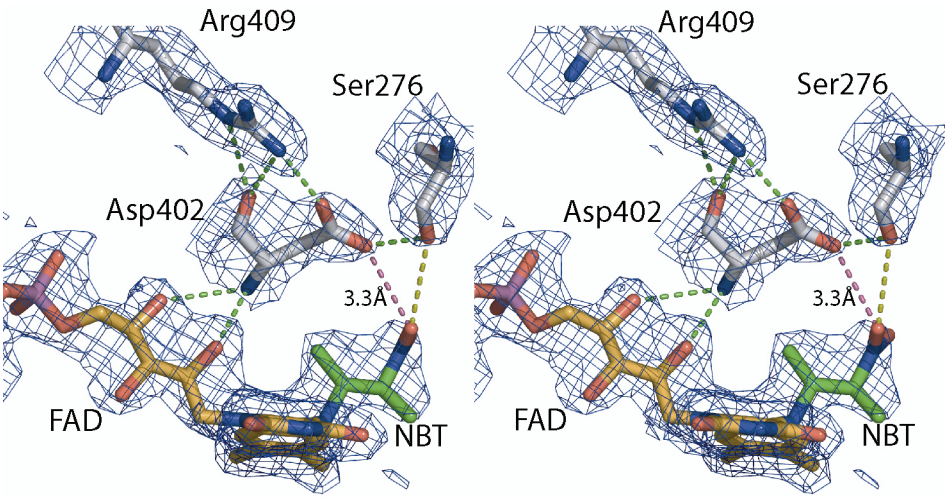


Figure 5.11 A divergent stereo view of ES* (subunit A) with the active site residues. The 2mF_o - DF_c maps shown in the figure are contoured at 1 σ level. Note the side chain conformation of Asp402 and Ser276.

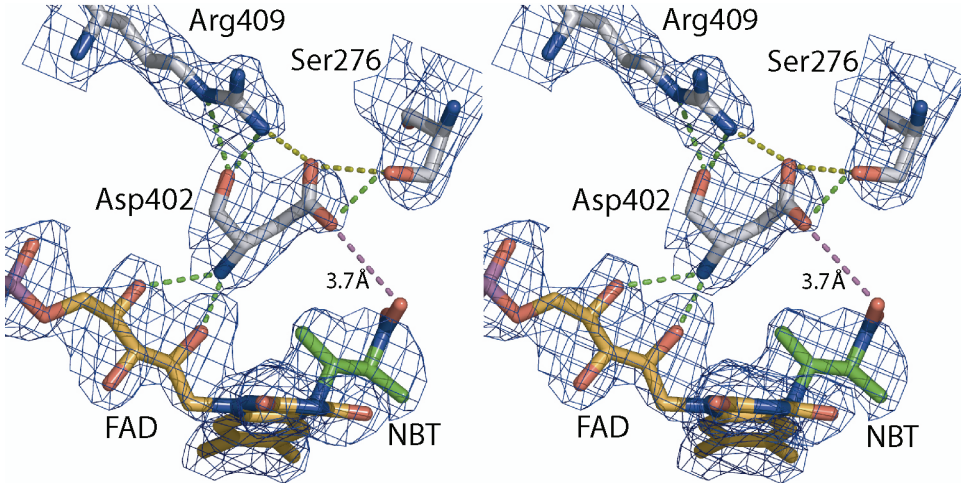


Figure 5.12 A divergent stereo view of ES* (subunit B) with the active site residues. The 2mF_o - DF_c maps shown in the figure are contoured at 1 σ level. Note the side chain conformation of Asp402 and Ser276.

Previous results have shown that the modified FAD adduct isolated from NAO is not stable in solution at high pH and converts rapidly to oxidized flavin (Gadda *et al.*, 1997). In contrast, our structural analysis of ES* suggests that stability of the adduct in the protein is likely the result of several factors (Figure 5.13) As suggested in the NAO reaction scheme (Figure 5.14), the decay of the adduct via step 9 may be related to the stability of the delocalized negative charge on N1 and O2 atoms of the pyrimidine ring. Our structures indicate that both the O2 and the N1 atoms form stable hydrogen bond interactions with two main chain amide protons and the hydroxyl group of Ser134, respectively (Figure 5.13). Some of the active sites of the ES* show that the nitro group of the adduct interacts with the side chain of Asp402, which is also coupled with Arg409. These interactions may reduce the ability to generate the NO_2^- leaving group as dictated by step 9 of scheme 1. In addition, the lack of O_2 reactivity of the reduced FAD-N5 adduct can be rationalized in part, by steric effects of the protein backbone on the *si*- face of the flavin (Figure 5.15). For example, the N5-(2-nitrobutyl) moiety forms on the *re*-face of the flavin and residues Trp169, Pro170 and Ser171 all pack the FAD *si*-face. Although there is a solvent excluded cavity on the *si*-face under the dimethylbenzene ring, access to C4a is not possible without a conformational change. Therefore upon release of the products, the *re*-face is most likely the site of O_2 binding in the oxidative half-reaction.

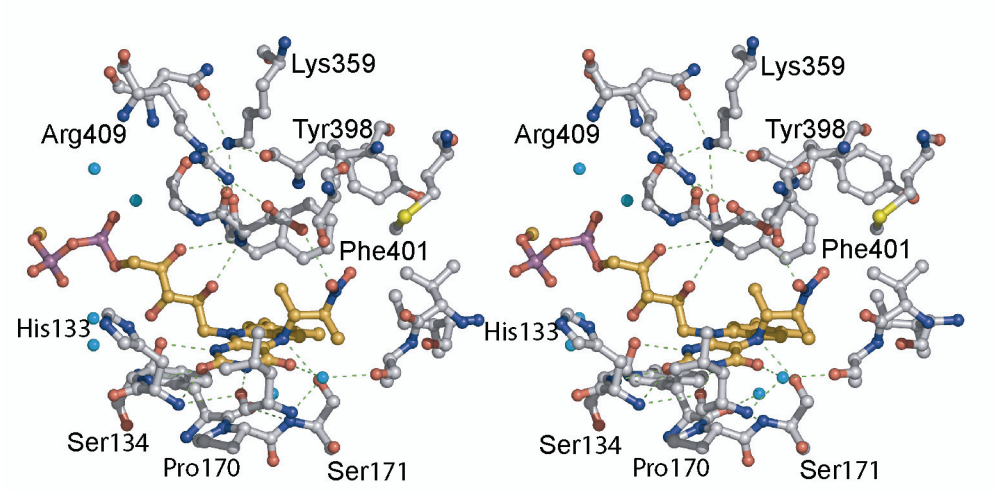


Figure 5.13 Divergent stereo image showing the environment around, and various hydrogen bonds stabilizing, the N5-(2-nitrobutyl)-1,5-dihydro-FAD moiety. Leu99, Phe273 although within 5 Å of nitrobutyl moiety have been removed to have a clear view. Residues not labeled in figure are: Ile92, Val95, Ala96, Leu131, Trp169, Ser276, V280, Met 283, Asp402, Gly403, Asn405.

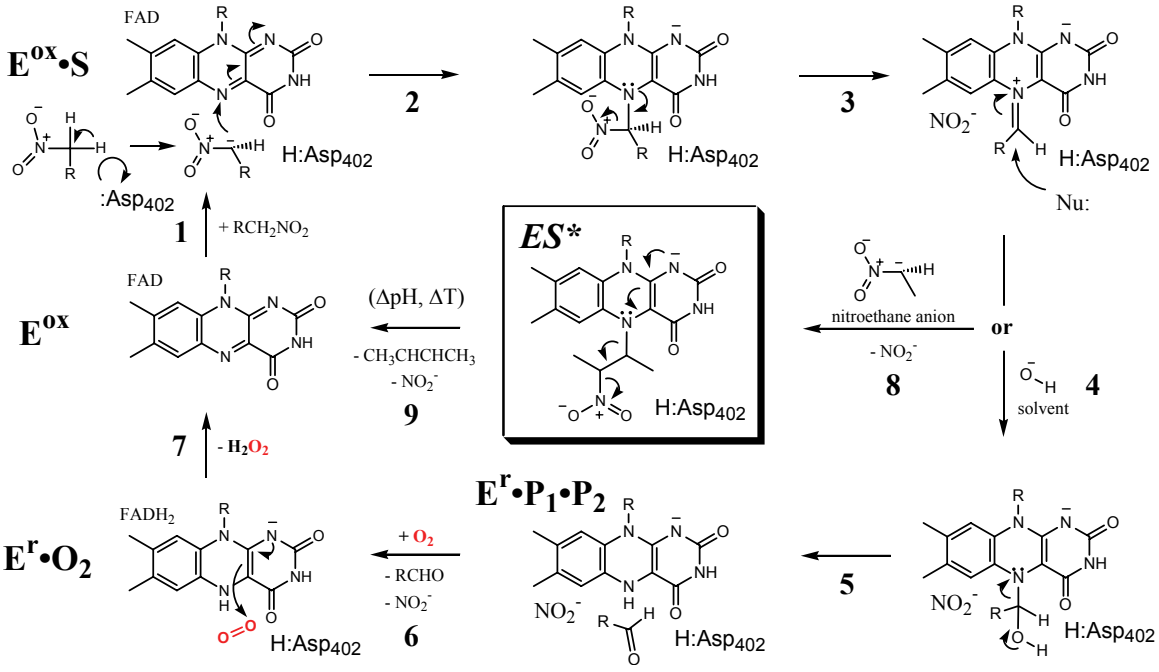


Figure 5.14 The proposed NAO reaction mechanism with nitroethane as substrate (Gadda & Fitzpatrick, 2000e)

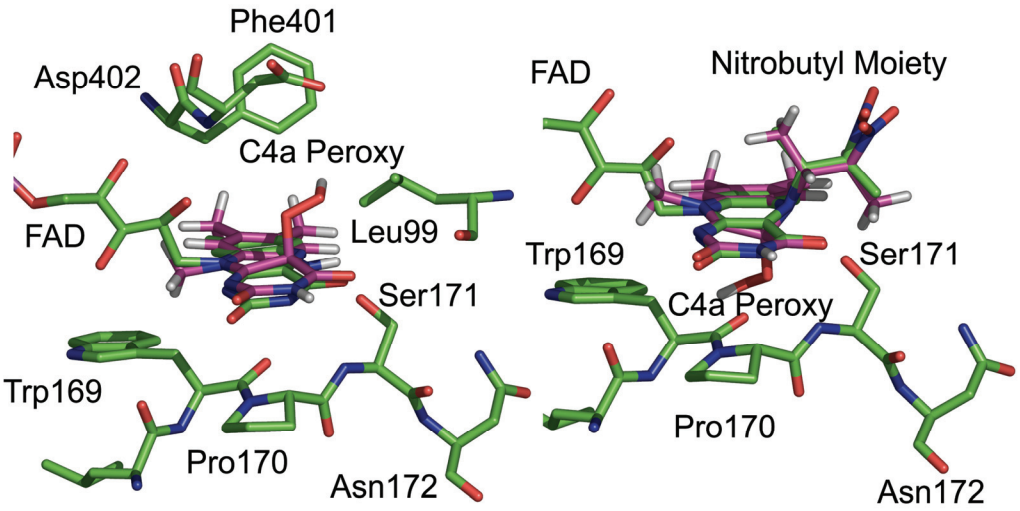


Figure 5.15 Hypothetical FAD-C4a:H_2O_2 complexes modeled with MOPAC simulations in *CHEM3D* using either a *re*-face complex (left, normal isoalloxazine ring system) or a *si*-face complex (right, N5-blocked system). Each model was overlapped with the isoalloxazine ring in either the E^{ox} or the ES* complex, respectively. Note that the *re*-face complex is not sterically hindered and that Arg409-Asp402 residue appear to be positioned to donate one or more protons in the oxidative half-reaction. In contrast, the *si*-face complex has several steric clashes and no proton donors nearby.

5.4 Comparison to the ACAD Family Members

The structural homologs of NAO include family members of the mitochondrial acyl-CoA dehydrogenases (ACAD) (Kim & Miura, 2004), acyl-CoA oxidase (ACO) I from *Arabidopsis thaliana* (Pedersen & Henriksen, 2005), ACO II from rat liver peroxisomes (Fitzpatrick *et al.*, 2005), and the acyl-ACP dehydrogenase from polyketide biosynthetic pathways (Watanabe *et al.*, 2003). Both ACADs and ACOs catalyze the oxidation of acyl-CoA thioesters to their corresponding *trans*-Δ2-enoyl-CoA. However, the reactions differ in the oxidative half-reactions. The ACAD family members transfer reducing equivalents to the electron transfer flavoprotein for use in the oxidative

phosphorylation cascades. In contrast, the ACOs are located in the peroxisome and use O_2 in the oxidative half-reaction to yield H_2O_2. NAO shares more sequence similarity (41 - 45%) with various ACAD family members and only 7 - 9% sequence similarity with the two ACOs. The regions of homology are not limited to particular domains and are spread throughout the entire primary sequence, with the exception of the C-terminal region of the ACOs (Figure 5.3). The overall structural similarity among the family members is also evident from the r.m.s. differences (Table 5.2) that range from 1.7 Å for the medium chain acyl-CoA dehydrogenase (MCAD) to 2.6 Å for the ACO-I (Figure 5.16). The *Q* scores range between 0.64 (MCADs) and 0.2 (ACOs) (Krissinel & Henrick, 2004). Despite these similarities, numerous molecular replacement trials failed to solve the structure of the ES* complex of NAO.

Table 5.2 Similarity between NAO-ES* and selected homologous proteins

Homolog	Q-score[a]	R.M.S.D.	Number C_α
MCAD[b]	0.63	1.67	369
ACO – I[c]	0.22	2.59	328
ACO – II[d]	0.21	3.11	331
BSCAD[e]	0.55	1.95	352
Rat SCAD[f]	0.57	2.01	355
ICAD	0.57	1.90	359
Acyl-ACPD[h]	0.48	2.15	327
Glutaryl-CoAD[i]	0.50	2.07	349

[a] Q score is defined as $N_{algn} * N_{algn} / (1+(RMSD/R_0)^2) * N_{res1} * N_{res2}$, where N_{algn} is the length of alignment and N_{res1} and N_{res2} are the number of total residues in the NAO and the target PDB structure. Calculations were done using the *Secondary Structure Matching* server (Krissinel & Henrick, 2004). [b] Medium chain acyl-CoA dehydrogenase from PDB entry (3mde). [c] Acyl-CoA Oxidase isoforms I from PDB entry (1w07). [d] Acyl-CoA Oxidase II from PDB entry (1is2). [e] Bacterial acyl-CoA dehydrogenase from PDB entry (1buc). [f] Short chain acyl-CoA dehydrogenase from PDB entry (1jqi). [g] Isovaleryl-CoA dehydrogenase from PDB entry (1ivh). [h] Acyl-carrier protein dehydrogenase from PDB entry (1r2j). [i] Glutaryl-CoA dehydrogenase decarboxylase from PDB code (1siq).

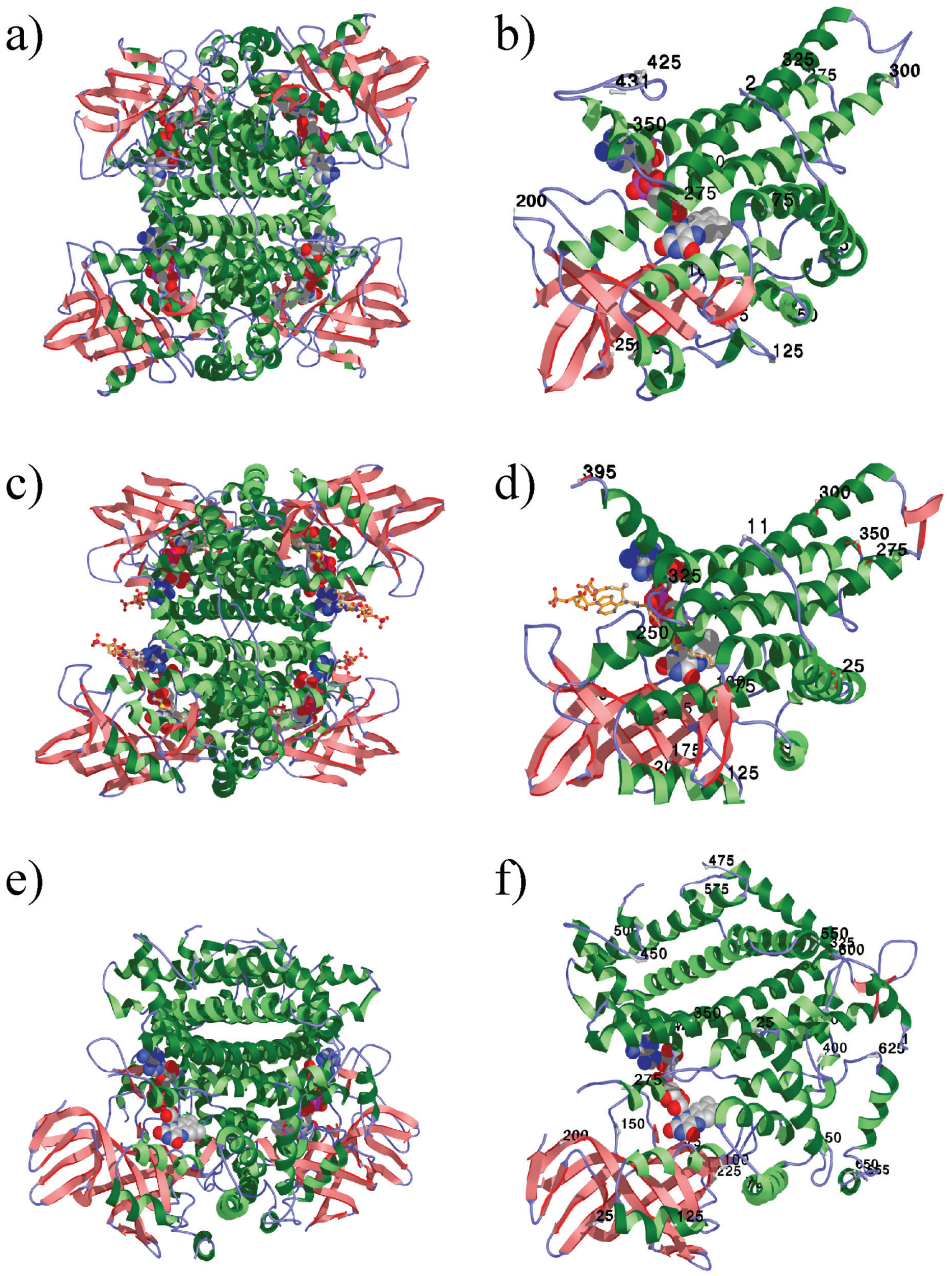


Figure 5.16 Crystal structures of nitroalkane oxidase, medium chain acyl-CoA dehydrogenase, and acyl-CoA oxidase displayed in similar orientations and scales. Shown are the Cα ribbon traces for: a) the α_4 holoenzyme of NAO; b) one subunit of NAO with FAD shown with space filled, CPK colored atoms; c) the α_4 holoenzyme of pig liver mitochondrial medium-chain acyl-CoA dehydrogenase (from PDB code 3mde, (Kim, J.-J. P. *et al.*, 1993)); d) one subunit of MCAD with FAD and with octanoyl-CoA in ball and stick format; e) the α_2 holoenzyme of rat liver peroxisomal acyl-CoA oxidase-II holoenzyme (from PDB code 1is2, (Nakajima *et al.*, 2002)); f) one subunit of ACO with FAD.

The structural, mechanistic, mutagenic and spectroscopic analysis of the medium chain ACAD (MCAD) provides exquisitely detailed proposals for the structural contributions of active site residues to catalysis (Ghisla & Thorpe, 2004; Kim & Miura, 2004). For example, the structures of MCAD bound to transition-state analogs, substrates or products clearly show that the acyl-CoA moiety occupy the N1 channel adjacent to the ribityl portion of FAD (Kim, J. J. *et al.*, 1993; Satoh *et al.*, 2003) (Figure 5.17 *c*). In contrast, our structure of EI shows that spermine occupies the N5 channel on the opposite side of the FAD (Figure 5.17 *a*). Moreover, the substrates cannot bind productively to NAO via the N1 channel since it is interrupted by the side chains of Phe273, Leu408 and Arg415. NAO does not react with acyl-CoA substrates, perhaps because several important binding determinants in the N1 channel of the ACADs are missing in NAO. For example, several residues that form hydrogen bonds with acyl-CoA substrates in the N1 channel of MCAD are replaced by alanines in NAO (Figure 5.17 *c, d*). Similarly, substrates cannot access the MCAD active site via the analogous N5 channel because the entrance of the channel is blocked by the side chain of His65. Moreover, the N5 channel in MCAD is interrupted by the side chains of Thr255 and Thr102, which results in two isolated cavities (Figure 5.17 *b*).

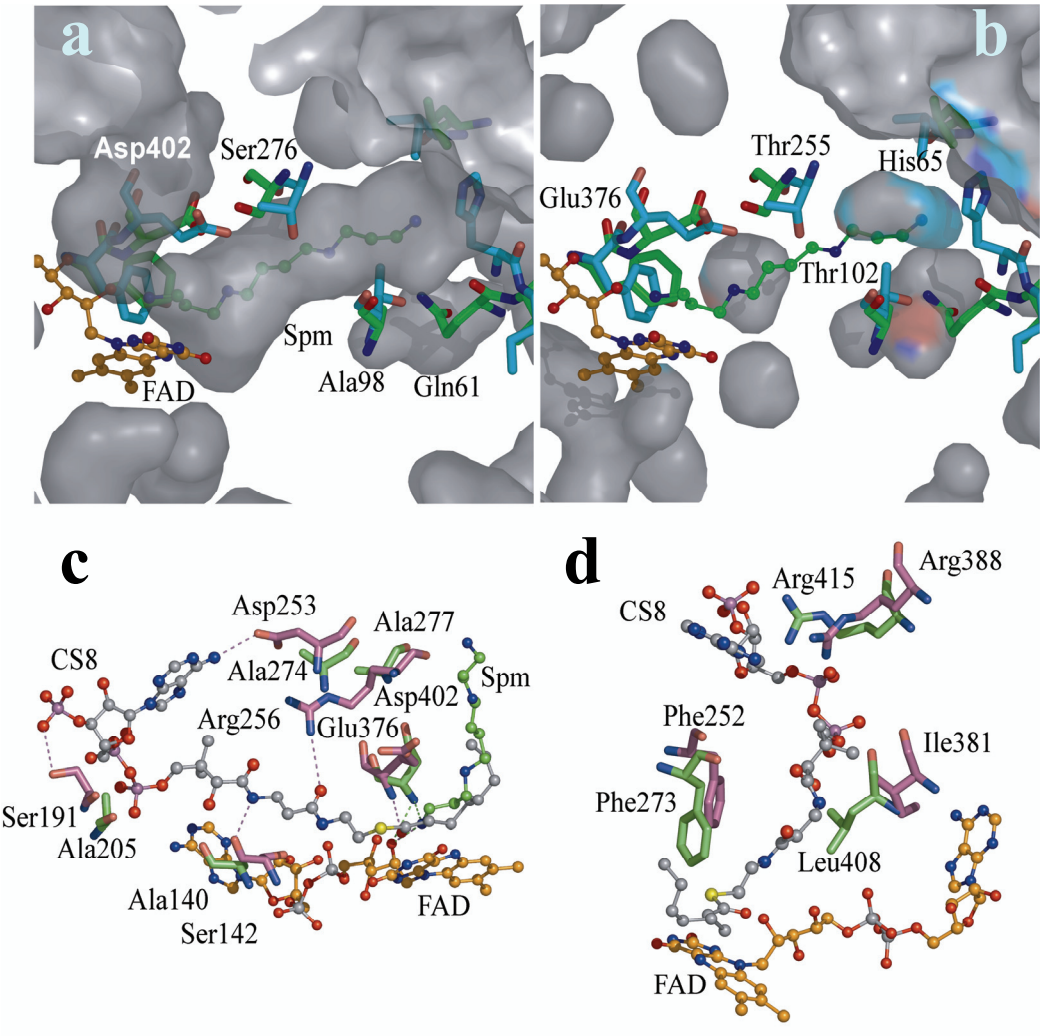


Figure 5.17 a) The N5 substrate access channel in NAO along with the active site NAO residues, inhibitor spermine (green) and the FAD (orange) overlapped with the corresponding active site residues from MCAD transition state complex (PDB code, 1udy) (cyan).

b) The N5 buried cavities in MCAD transition state complex (PDB code, 1udy) complex overlapped with corresponding active site residues and FAD from NAO EI complex. The MCAD and NAO residues are labeled and colored as in (a). The cavities in the MCAD are interrupted by Thr255 and blocked from the exterior by His65.

c & d) MCAD residues lining the N1 channel (magenta) and the corresponding residues from NAO (green). The residues in N1 channel of NAO sterically block the access to active site FAD. The hydrogen bond donors and acceptors for the substrate positioning and binding are missing in NAO in contrast to MCAD.

Apart from the differences in the substrate access channel, the structural comparisons of NAO with ACADs also highlight the differences in the hydrogen bonding associated with the active site base, FAD and substrates. All these interactions are important determinants in guiding a particular reaction chemistry. The active site in NAO can be described as a two layer sandwich. In contrast, the ACADs, has three layers in the active site, the FAD, the substrate and the active site loop. Most of the hydrogen bonding interactions between isoalloxazine ring and the protein backbone are conserved in NAO and are similar to those of the ACADs (Figure 5.18). The hydrogen bond between 2'OH of the ribityl chain and the terminal nitrogen of spermine is also similar to the one observed between the carbonyl thioester of acyl-CoA substrate and 2'OH of the ribityl chain of FAD in ACADs. This particular hydrogen bond has been shown to lower the pKa of the acyl-CoA substrates in ACADs and is suggested to play a similar role in NAO. In NAO, 2'OH and 4'OH of ribityl chain of the FAD is hydrogen bonded to the amide nitrogen of active site base, Asp 402. The hydrogen bond interaction between Asp 402 and 2'OH and 4'OH of the ribityl chain are unique to NAO and not present in other ACAD superfamily members. In addition to above mentioned hydrogen bonds, 4'OH of the ribityl chain in NAO also forms hydrogen bonds with other residues on the active site loop (amide nitrogen of Gly403 and carbonyl oxygen of Leu400). These hydrogen bonds are missing in MCADs.

In MCAD, all the oxygens of the FAD phosphates face towards *si*-face leaving more room for the acyl-CoA substrates on the *re*-face. For example, in MCAD, the two oxygens of the first phosphate hydrogen bond with NH1 of Arg281 and amide nitrogen of

Gly353 from the adjacent subunit whereas in NAO the oxygens from the first phosphate face towards *re*-face and form hydrogen bond with amide N of Gly404 from the same subunit and Met379 from the other subunit. Both in NAO and MCAD the oxygens from the second phosphate face the *si*-face of FAD and are stabilized through various hydrogen bonds. In NAO, one of the oxygens from the second phosphate hydrogen bonds with the NH1 of Arg304 from the other subunit, which suggests that dimer formation is not only structurally important in NAO but is also critical part of FAD environment.

The Arg409 forms an ionic and hydrogen bond interaction with Asp402 in our NAO structures. This feature is unique to NAO compared to the other homologs. Interestingly, the guanidine group of Arg256 in MCAD occupies an approximately analogous position. However, the group does not from a hydrogen bond to the analogous active site base, Glu376, but rather packs the hydrophobic portion against the acidic side chain. Thus a positive charge is approximately conserved in the three dimensional space, but is donated from residues on opposite sides of the active site cavity. In conclusion, the substrate specificity of NAO can be attributed to the differences in substrate access channel, FAD environment and hydrogen bonding interactions with the active site base.

The solvent accessible surface area (SASA) of N5 (~6.0 $Å^2$) in E^{ox} is comparable to ACOs (~5.5 $Å^2$), but far higher than the observed values for MCAD (2.4 $Å^2$). In contrast, the SASA value for C4a is lowest for NAO (1.2 $Å^2$) in comparison to values reported for other homologs (ACOs -2.6 $Å^2$ and MCAD – 1.9 $Å^2$). The substrate binding and adduct formation in NAO completely shield C4a and N5 from solvent, as evident

from zero values of SAS. This observation suggests that products of NAO reaction cycle likely dissociate in order to make the C4a accessible to O_2. This is in contrast to acyl-CoA products, which bind tightly in the active site of ACADs and thus prevent the entry of O_2 in active site (Kim & Miura, 2004).

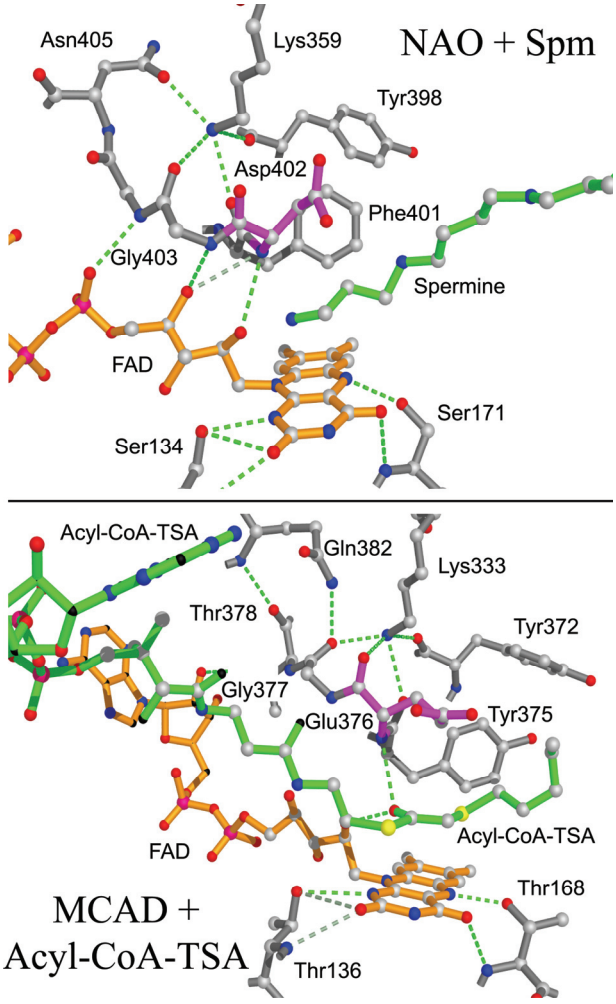


Figure 5.18 Active site comparison of the EI complex of NAO and the MCAD:acyl-CoA-TSA complex (PDB code 1udy)

Another striking feature of the ACAD family members is the presence of a structurally conserved Lys, which corresponds to residue 359 in NAO. The charged N^{ε} atom of Lys is surrounded by four oxygen atoms similar to the K^+ selectivity filter in the K^+ channel (Figure 5.6) (Doyle *et al.*, 1998). In all the ACAD superfamily members, the site appears to function to interrupt the "J" α-helix and promote the distinct turn in the peptide backbone trace. This consequently places the active site base Asp402 and the FAD in proper position for reaction (Figure 5.18,). The Lys is not conserved in the primary amino acid sequence of iso(3)valeryl-CoA dehydrogenase (ICAD) (Figure 5.6) (Tiffany *et al.*, 1997) but a structurally analogous Arg381 is present. Although the active site base, Glu254 in ICAD resides on α helix "G" rather than on the J-K loop (Figure 5.3), a structurally conserved positive charge is still required for twisting the helix K in the active site. A helical twist of 40^o between helix J and helix K in ICADs is very similar to the one observed in other ACADs (helical twist 45^o). The turning of K α helix is critical in ACADs to maintain the active site architecture. It allows the acyl-CoA substrates, which bind in the N1 substrate channel to access the active site FAD. Although, NAO substrates bind via N5 substrate access channel, a conserved helical twist of 40^o between helix J and helix K is still observed and is suggested to be important for the proper positioning of the active site base Asp402.

6 CONCLUSIONS

Nitroalkane oxidase (NAO), a flavoenzyme, from *Fusarium oxysporum* catalyzes the oxidation of the neutral nitroalkanes to the corresponding aldehydes or ketones with production of H_2O_2 and nitrite. Although cDNA sequence analysis shows that the flavoenzyme is a new member of the acyl-CoA dehydrogenase (ACAD) superfamily. NAO does not bind or turnover acyl-CoA substrates. In contrast to ACADs, NAO also follows a different reaction mechanism and forms a stable N5 covalent adduct during the catalytic cycle. Biochemical modifications of NAO suggested the presence of Cys397, Tyr398, Asp402 and Arg in the active site but the specific role of these residues in the active site could not be determined because of the lack of three dimensional structure of the enzyme (Gadda *et al.*, 2000a; Gadda *et al.*, 2000b; Gadda *et al.*, 2001; Valley & Fitzpatrick, 2003a). Crystal structure studies of NAO in different reaction states provide insights into the active site architecture, substrate binding channel, and possible reasons for relative stability of N5-FAD adduct in the enzyme versus the solution state.

The bottlenecks in solving the crystal structures of NAO were several-fold and have been described and discussed throughout this thesis. Conventionally, in protein crystallography, getting good diffraction quality crystals is a major impediment; but in the case of NAO, crystals were obtained in a relatively short period of time. NAO was crystallized in four different space groups, but the highest diffraction quality crystals had an exceptionally long cell edge. In addition to the long cell edge, phasing of the protein was a challenging crystallographic problem and took significant computational efforts.

The structure of ES*, the N5 covalent adduct of NAO, was solved to 2.2 Å resolution using phases calculated from location of 52 SeMet sites. To our knowledge, this is the first crystal structure of a flavoprotein trapped during turnover of authentic substrates as a covalent adduct at N5-FAD. We also solved the structure of oxidized NAO to 2.07 Å resolution. The asymmetric unit has 1.5 homotetramers out of which only four active sites are occupied with a weak inhibitor, spermine. Together, these structures provide insights into: a) substrate preference and access to the FAD, b) the reaction mechanism, and c) the features that differentiate NAO from the structural but not the functional homologs in the ACAD family.

The structure of the ES* complex shows clearly resolved electron density for the 2-nitrobutyl moiety at the N5 position of the isoalloxazine ring. The observation of an N5- adduct also supports the carbanion reaction mechanism for NAO. The isoalloxazine ring of the flavin forms numerous hydrogen bonding interactions, which are suggested to play an important role in adduct stabilization in the active site of the protein. A delocalized negative charge between N1 and O2 positions of the isoalloxazine ring is stabilized via hydrogen bonding interactions with the side chain of Ser134 and amide protons on the nitrogen atoms of Ser134 and His133, respectively. The nitro group of the adduct forms hydrogen bond with OD2 of the Asp402. Asp402 further forms hydrogen bond and ionic interaction with Arg409. The N3 and N5 of the isoalloxazine ring are hydrogen bonded to O of Leu131 and OG of Ser171, respectively.

The structure of the EI complex of NAO reveals a substrate access channel that terminates at the N5 atom of the flavin. In ACADs, the N5 cavity is not accessible from

the exterior of the protein and the substrates bind to the active site via N1 substrate access channel. Nitroalkane oxidase, in contrast to its ACAD homologs, has sterically occluded N1 channel so that the substrates access the active site via N5 channel. The N1 cavity in NAO also has missing hydrogen bond donors and acceptors, which are critical for proper positioning of the acyl-CoA substrates in ACADs. Although NAO and ACADs belong to the same superfamily, the differences in the substrate channel explain the observation that they do not cross react. The terminal nitrogen of the spermine forms a hydrogen bond with the 2'OH of the ribityl chain of FAD, which is similar to the hydrogen bond between the carbonyl oxygen of the thioester substrate and 2' OH of the flavin ribityl chain in case of ACADs. This particular hydrogen bond is critical for lowering the pKa of the substrates in ACADs and is suggested to have a similar role in NAO. Asp402 is the active site base in NAO, whereas in ACADs the analog is Glu376. The electrostatic interactions between Lys359 and the residues of active site J-K loop are conserved in all the members of ACAD superfamily, except for iso-valeryl acyl-CoA dehydrogenase (ICAD). In ICAD, a similar role is played by Arg381 despite the fact that the active site base is not located on the J-K loop. The crystal structural studies of NAO suggest that NAO and ACADs diverged sufficiently from a common ancestor to support different substrate specificity and reaction mechanism.

The structural overlay of ES* and EI revealed significant movement in the side chain of Asp402 (1.26 Å). The side chain movement is suggested to help in α-proton abstraction from the substrate. The side chain conformation of Asp402 in NAO subunits is coupled to orientation of Ser276 side chain. Most of the large r.m.s.d. shifts between

the E^{ox}, EI, and ES* crystal structures are observed for the residues, which are far away from the active site and are located on the protein surface. For example, the protein backbone and side chains of Glu 342 and Glu344, which are present at the mouth of N5 channel, have main chain r.m.s.d. of ~0.76 Å and side chain r.m.s.d. of ~2.2 Å.

The crystal structures of EI and ES* reveal that the protein backbone packs tightly against the *si*-face of the flavin in NAO. Thus, the formation of C4a-peroxy adduct of FAD is sterically occluded on the *si*-face. The structural overlap of the hypothetical model of C4a-hydroperoxide adduct of FAD and the active site FAD in NAO shows that *re*-face of the flavin has sufficient volume to allow C4a peroxy formation. Since the N5 position is blocked in ES* and *si*-face is sterically occluded for C4a peroxy formation, the adduct exhibits resistance towards O_2 reactivity. To conclude, NAO crystal structures have helped in defining a structural basis for the observed biochemical properties of NAO. The knowledge of structures will also be useful in designing the future studies on this new member of ACAD superfamily.

7 FUTURE DIRECTIONS

The goal of this thesis was to solve the three dimensional structure of NAO and its different reaction intermediates. The structural comparison of E^{ox}, EI and ES* reveal significant differences in the conformation of the active site residues, which suggest their role in substrate specificity, reaction mechanism and N5-adduct stabilization. The specific functional role of these residues needs to be further evaluated through mutagenesis and crystallographic studies.

Nitroalkane oxidase is a unique flavoenzyme in many aspects. One of the characteristic features of the resting state of the enzyme is its selectivity towards neutral nitroalkane. In contrast, the cationic imine formed during the catalytic cycle does react with nitronates to form stable N5-adduct (Valley et al., 2005). The kinetic analysis of D402A mutant shows that enzyme activity can be rescued with the deprotonated substrate (Valley & Fitzpatrick, 2003a). This suggests that a negative charge on Asp402 in the resting state of the enzyme could inhibit the nitronates from binding the active site. It will be interesting to study the crystal structures of D402A mutant in complex with NAO inhibitors like benzoate or valarate to determine whether negatively charged substrates still bind via N5 channel or if they bind in N1 channel that has a positively charged guanidinio group from Arg409. Although the N1 channel in NAO has the largest volume, the active site access through this channel is sterically restricted because of Phe273. Different mutants of Phe273 will help in studying substrate binding possibilities through the N1 channel. Site directed mutagenesis and crystallographic analysis of Leu99, Leu131, His133, Ser134, Trp169, Ser171, Phe273, Asp402, Arg409 mutants will shed

more light on how the active site environment might help in the stabilization of N5-adduct.

The crystal structures of NAO also form the basis for directed evolution studies. Although the overall protein fold of NAO is similar to ACADs, there are differences in the active site residues and in the residues lining the substrate channel. This observation along with cDNA analysis, which identified NAO as a member of ACAD superfamily (Daubner et al., 2002), suggest that NAO and ACADs have diverged from a common ancestor to carry out different reactions. Structural overlays of NAO with ACAD homologs suggest that residues like Gln61, Ala98, Ala140, Ala205, Phe273, Ala274, Ser276, Ala277, Asp 402, Leu408, and Arg415 are likely to define NAO substrate specificity. Kinetic studies of these mutants can help in studying the natural process of evolution at the laboratory level. It will be also interesting to see if any of these mutants will act as NAO-ACAD chimeras and strengthen the evolutionary link between NAO and other ACAD homologs.

Future studies could also be focused on understanding the role of FAD in the NAO reaction mechanism. Flavins are versatile redox centers, which depending on their redox potential, will either accept or donate electrons. The redox potential of the flavins is largely determined by the protein environment (Massey, 2000). In different structures of NAO, we observe large shifts in the side chain conformations of Asn172 and Asn405. It is interesting to know that active site asparagines are the characteristic feature of flavoproteins like flavodoxin from *Helicobacter pylori* (Freigang *et al*., 2002),

cholesterol oxidase (Sampson, 2001) and choline oxidase (Lountos *et al*., unpublished work) etc. The Asn mutants of flavodoxins and cholesterol oxidase have suggested that asparagines play an important role in controlling the redox potential of flavins (Yin *et al.*, 2001; Carr *et al.*, 1990). Although Asn172 and Asn405 in NAO are about ~6 Å away from the isoalloxazine ring, single mutant studies will provide insights into their potential role in NAO. The knowledge of the active site residues and their interaction with FAD can also be used to modify NAO chemistry. For example, in the case of NAO, the ribityl chain of the FAD forms some unique hydrogen bond interactions that are not present in other homologs. The differences in the hydrogen bonding pattern is suggested to play an important role in defining the active site environment of NAO. The role of particular hydrogen bond interactions in the reaction mechanism can be evaluated by use of deoxy-flavins.

The scope of the NAO crystal structures is not limited to protein crystallography or protein biochemistry, but the structural knowledge can also be applied in designing better biocatalysts for bioremediation, in making biosensors that can sense different nitro species based on their charge, and in other bioelectronic applications.

REFERENCES

Abrahams, J. P. and Leslie, A. G. W. (1996). Methods used in the structure determination of bovine mitochondrial F_1 ATPase. *Acta Cryst., D52*, 30-42.

Adams, P. D., Pannu, N. S., Read, R. J. and Brunger, A. T. (1999). Extending the limits of molecular replacement through combined simulated annealing and maximum-likelihood refinement. *Acta Cryst., D55* (Pt 1), 181-190.

Ahlner, J., Andersson, R. G., Torfgard, K. and Axelsson, K. L. (1991). Organic nitrate esters: clinical use and mechanisms of actions. *Pharmacol. Rev., 43* (3), 351-423.

Alston, T. A., Mela, L. and Bright, H. J. (1977). 3-Nitropropionate, the toxic substance of Indigofera, is a suicide inactivator of succinate dehydrogenase. *Proc. Natl. Acad. Sci. U. S. A., 74*, 3767–3771.

Alston, T. A., Porter, D. J. and Bright, H. J. (1983). Suicide inactivation of D-amino acid oxidase by 1-chloro-1-nitroethane. *J. Biol. Chem., 258*, 1136-1141.

Aulak, K. S., Koeck, T., Crabb, J. W. and Stuehr, D. J. (2004). Dynamics of protein nitration in cells and mitochondria. *Am. J. Physiol. Heart Circ. Physiol., 286* (1), H30-38.

Aulak, K. S., Koeck, T., Crabb, J. W. and Stuehr, D. J. (2004). Proteomic method for identification of tyrosine-nitrated proteins. *Methods Mol. Biol., 279*, 151-165.

Aulak, K. S., Miyagi, M., Yan, L., West, K. A., Massillon, D., Crabb, J. W. and Stuehr, D. J. (2001). Proteomic method identifies proteins nitrated in vivo during inflammatory challenge. *Proc. Natl. Acad. Sci. U. S. A., 98* (21), 12056-12061.

Bach, R. D. and Dmitrenko, O. (2003). Electronic requirements for oxygen atom transfer from alkyl hydroperoxides. Model studies on multisubstrate flavin-containing monooxygenases. *J. Phys. Chem., 107* (46), 12851-12861.

Bach, R. D. and Dmitrenko, O. (2004). Model studies on p-hydroxybenzoate hydroxylase. The catalytic role of Arg-214 and Tyr-201 in the hydroxylation step. *J. Am. Chem. Soc., 126* (1), 127-142.

Bach, R. D., Thorpe, C. and Dmitrenko, O. (2002). C-H...carboxylate oxygen hydrogen bonding in substrate activation by acyl-CoA dehydrogenases: Synergy between the H-bonds. *J. Phys. Chem., 106* (16), 4325-4335.

Barna, T. M., Khan, H., Bruce, N. C., Barsukov, I., Scrutton, N. S. and Moody, P. C. (2001). Crystal structure of pentaerythritol tetranitrate reductase: "flipped" binding geometries for steroid substrates in different redox states of the enzyme. *J. Mol. Biol., 310* (2), 433-447.

Barton, G. J. (1993). ALSCRIPT - A Tool to Format Multiple Sequence Alignments. *Prot. Eng., 6*, 37-40.

Battaile, K. P., Molin-Case, J., Paschke, R., Wang, M., Bennett, D., Vockley, J. and Kim, J. J. (2002). Crystal structure of rat short chain acyl-CoA dehydrogenase complexed with acetoacetyl-CoA: comparison with other acyl-CoA dehydrogenases. *J. Biol. Chem., 277* (14), 12200-12207.

Battaile, K. P., Nguyen, T. V., Vockley, J. and Kim, J. J. (2004). Structures of isobutyryl-CoA dehydrogenase and enzyme-product complex: comparison with isovaleryl- and short-chain acyl-CoA dehydrogenases. *J. Biol. Chem., 279* (16), 16526-16534.

Beinert, H. and Page, E. (1957). On the mechanism of dehydrogenation of fatty acyl derivatives of coenzyme A. Oxidation-reductions of the flavoproteins. *J. Biol. Chem., 225* (1), 479-497.

Bell, R. P. (1974). Recent advances in the study of kinetic hydrogen isotope effects. *Chem. Soc. Rev., 3* (4), 513 –544.

Bhushan, B., Halasz, A., Spain, J. C. and Hawari, J. (2004). Initial reaction(s) in biotransformation of CL-20 is catalyzed by salicylate 1-monooxygenase from *Pseudomonas* sp. strain ATCC 29352. *Appl. Environ. Microbiol., 70* (7), 4040-4047.

Bhushan, B., Trott, S., Spain, J. C., Halasz, A., Paquet, L. and Hawari, J. (2003). Biotransformation of Hexahydro-1,3,5-Trinitro-1,3,5-Triazine (RDX) by a Rabbit Liver Cytochrome P450: Insight into the Mechanism of RDX Biodegradation by *Rhodococcus* sp. Strain DN22. *Appl. Environ. Microbiol., 69* (3), 1347-1351.

Binda, C., Angelini, R., Federico, R., Ascenzi, P. and Mattevi, A. (2001). Structural bases for inhibitor binding and catalysis in polyamine oxidase. *Biochemistry, 40* (9), 2766-2776.

Binda, C., Li, M., Hubalek, F., Restelli, N., Edmondson, D. E. and Mattevi, A. (2003). Insights into the mode of inhibition of human mitochondrial monoamine oxidase B from high-resolution crystal structures. *Proc. Natl. Acad. Sci. U. S. A., 100* (17), 9750-9755.

Binda, C., Mattevi, A. and Edmondson, D. E. (2002a). Structure-function relationships in flavoenzyme-dependent amine oxidations: a comparison of polyamine oxidase and monoamine oxidase. *J. Biol. Chem., 277*, 23973-23976.

Binda, C., Newton-Vinson, P., Hubalek, F., Edmondson, D. E., and and Mattevi, A. (2002b). Structure of human monoamine oxidase B, a drug target for the treatment of neurological disorders. *Nat. Struct. Biol., 9*, 22-26.

Binks, P. R., French, C. E., Nicklin, S. and Bruce, N. C. (1996). Degradation of pentaerythritol tetranitrate by *Enterobacter cloacae* PB2. *Appl. Environ. Microbiol., 62* (4), 1214-1219.

Blehert, D. S., Fox, B. G. and Chambliss, G. H. (1999). Cloning and sequence analysis of two *Pseudomonas* flavoprotein xenobiotic reductases. *J. Bacteriol., 181* (20), 6254-6263.

Blehert, D. S., Knoke, K. L., Fox, B. G. and Chambliss, G. H. (1997). Regioselectivity of nitroglycerin denitration by flavoprotein nitroester reductases purified from two *Pseudomonas* species. *J. Bacteriol., 179* (22), 6912-6920.

Blundell, T. N. and Johnson, L. N. (1976). *Protein Crystallography*. New York: Academic Press.

Bogdan, C. (2001). Nitric oxide and the immune response. *Nat. Immunol., 2* (10), 907-916.

Bordwell, F. G. and Boyle Jr., W. J. (1975). Kinetic isotope effects for nitroalkanes and their relation to transition-state structure in proton-transfer reactions. *J. Am. Chem. Soc., 97*, 3447 –3452.

Bottaro, J. C. (1996). Recent advances in explosives and solid propellants. *Chemistry & Industry (London), 7*, 249-252.

Bricogne, G. (1997). Bayesian statistical viewpoint on structure determination:basic concepts and examples. *Methods Enzymol., 276*, 361-423.

Bright, H. J. and Porter, D. J. T. (1975). Flavoprotein oxidases. In P. Boyer (Ed.), *The Enzymes* (3 ed., Vol. XII, pp. 421-505). New York: Academic Press.

Bruice, T. C. (1980). Mechanisms of flavin catalysis. *Acc. Chem. Res., 13*, 256 –262.

Brünger, A. T. (1992). The free R value: a novel statistical quantity forassessing the accuracy of crystal structures. *Nature, 355*, 472-474.

Brünger, A. T. (1997). Free R value: cross-validation in crystallography. *Methods Enzymol., 277*, 366-396.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. and Warren, G. L. (1998a). Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Cryst., D54* (Pt 5), 905-921.

Brünger, A. T., Adams, P. D. and Rice, L. M. (1998b). Recent developments for the efficient crystallographic refinement of macromolecular structures. *Curr. Opin. Str. Biol., 8*, 606–611.

Brünger, A. T., Krukowski, A. and Erickson, J. (1990). Slow-Cooling Protocols for Crystallographic Refinement by Simulated Annealing. *Acta Cryst., A46*, 585-593.

Butterfield, D. A. (2004). Proteomics: a new approach to investigate oxidative stress in Alzheimer's disease brain. *Brain Res., 1000* (1-2), 1-7.

Byron, C. M., Stankovich, M. T. and Husain, M. (1990). Spectral and electrochemical properties of glutaryl-CoA dehydrogenase from *Paracoccus denitrificans*. *Biochemistry, 29* (15), 3691-3700.

Cane, D. E., Walsh, C. T. and Khosla, C. (1998). Harnessing the biosynthetic code: combinations, permutations, and mutations. *Science, 282* (5386), 63-68.

Carr, M. C., Curley, G. P., Mayhew, S. G. and Voordouw, G. (1990). Effects of substituting asparagine for glycine-61 in flavodoxin from *Desulfovibrio vulgaris* (Hildenborough). *Biochem. Int., 20* (6), 1025-1032.

Carroll, B. J., Moss, S. J., Bai, L., Kato, Y., Toelzer, S., Yu, T. W. and Floss, H. G. (2002). Identification of a set of genes involved in the formation of the substrate for the incorporation of the unusual "glycolate" chain extension unit in ansamitocin biosynthesis. *J. Am. Chem. Soc., 124* (16), 4176-4177.

Chen, Z., Zhang, J. and Stamler, J. S. (2002). Identification of the enzymatic mechanism of nitroglycerin bioactivation. *Proc. Natl. Acad. Sci. U. S .A., 99* (12), 8306-8311.

Clark, N. G., Croshaw, B., Leggetter, B. E. and Spooner, D. F. (1974). Synthesis and antimicrobial activity of aliphatic nitro compounds. *J. Med. Chem., 17* (9), 977-981.

Coleman, N. V., Spain, J. C. and Duxbury, T. (2002). Evidence that RDX biodegradation by *Rhodococcus* strain DN22 is plasmid- borne and involves a cytochrome p-450. *J. Appl. Microbiol., 93* (3), 463-472.

Cowtan, K. and Main, P. (1996). Phase combination and cross validation in iterated density-modification calculations. *Acta Cryst., D52*, 43-48.

Crane, F. L. and Beinert, H. (1956b). On the mechanism of dehydrogenation of fatty acyl derivatives of coenzyme A. II. The electron-transferring flavoprotein. *J. Biol. Chem., 218* (2), 717-731.

Crane, F. L., Mii, S., Hauge, J. G., Green, D. E. and Beinert, H. (1956a). On the mechanism of dehydrogenation of fatty acyl derivatives of coenzyme A. I. The general fatty acyl coenzyme A dehydrogenase. *J. Biol. Chem., 218* (2), 701-706.

Crow, J. P. and Ischiropoulos, H. (1996). Detection and quantitation of nitrotyrosine residues in proteins: in vivo marker of peroxynitrite. *Methods Enzymol., 269*, 185-194.

Cruickshank, D. W. J. (1999). Remarks about protein structure precision. *Acta Cryst., D55*, 583-601.

Daubner, S. C., Gadda, G., Valley, M. P. and Fitzpatrick, P. F. (2002). Cloning of nitroalkane oxidase from Fusarium oxysporum identifies a new member of the acyl-CoA dehydrogenase superfamily. *Proc. Natl. Acad. Sci. U. S. A., 99* (5), 2702-2707.

de LaFortelle, E. and Bricogne, G. (1997). Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol., 276*, 472-494.

Dhawale, M. R., and and Hornemann, U. (1979). Nitroalkane oxidation by streptomycetes. *J. Bacteriol., 137* (2), 916-924.

Djordjevic, S., Pace, C. P., Stankovich, M. T. and Kim, J. J. (1995). Three-dimensional structure of butyryl-CoA dehydrogenase from *Megasphaera elsdenii*. *Biochemistry, 34* (7), 2163-2171.

Dmitrenko, O., Thorpe, C. and Bach, R. D. (2003). Effect of a charge-transfer interaction on the catalytic activity of Acyl-CoA dehydrogenase: A theoretical study of the role of oxidized flavin. *J. Phys. Chem., 107* (47), 13229-13236.

Doyle, D. A., Morais Cabral, J., Pfuetzner, R. A., Kuo, A., Gulbis, J. M., Cohen, S. L., Chait, B. T. and MacKinnon, R. (1998). The structure of the potassium channel: molecular basis of K^+ conduction and selectivity. *Science, 280* (5360), 69-77.

Drenth, J. (1999). *Principles of Protein X-ray Crystallography* (2 ed.).

Duncan, M. W. (2003). A review of approaches to the analysis of 3-nitrotyrosine. *Amino Acids, 25* (3-4), 351-361.

Ebert, S., Fischer, P. and Knackmuss, H. J. (2001). Converging catabolism of 2,4,6-trinitrophenol (picric acid) and 2,4-dinitrophenol by *Nocardioides simplex* FJ2-1A. *Biodegradation, 12* (5), 367-376.

Evans, G. and Pettifer, R. F. (2001). CHOOCH: a program for derivinf anomalous-scattering factors from X-ray fluorescence spectra. *J. Appl. Cryst., 34*, 82-86.

Evans, P. R. (1997). *Scaling of MAD Data.* Paper presented at the Proceedings of CCP4 Study Weekend.

Fink, C. W., Stankovich, M. T. and Soltysik, S. (1986). Oxidation-reduction potentials of butyryl-CoA dehydrogenase. *Biochemistry, 25* (21), 6637-6643.

Fitzpatrick, P. F., Orville, A. M., Nagpal, A. and Valley, M. P. (2005). Nitroalkane oxidase, a carbanion-forming flavoprotein homologous to acyl-CoA dehydrogenase. *Arch. Biochem. Biophys., 433* (1), 157-165.

Fraaije, M. W. and Mattevi, A. (2000). Flavoenzymes: diverse catalysts with recurrent features. *Trends Biochem. Sci., 25* (3), 126-132.

Francis, K., Russell, B. and Gadda, G. (2005). Involvement of a flavosemiquinone in the enzymatic oxidation of nitroalkanes catalyzed by 2-nitropropane dioxygenase. *J. Biol. Chem., 280* (7), 5195-5204.

Freeman, C. D., Klutman, N. E. and Lamp, K. C. (1997). Metronidazole. A therapeutic review and update. *Drugs, 54* (5), 679-708.

Freigang, J., Diederichs, K., Schafer, K. P., Welte, W. and Paul, R. (2002). Crystal structure of oxidized flavodoxin, an essential protein in *Helicobacter pylori*. *Protein Sci., 11* (2), 253-261.

French, C. E., Nicklin, S. and Bruce, N. C. (1998). Aerobic degradation of 2,4,6-trinitrotoluene by Enterobacter cloacae PB2 and by pentaerythritol tetranitrate reductase. *Appl. Environ. Microbiol., 64* (8), 2864-2868.

Gadda, G., Edmondson, R. D., Russell, D. H., and and Fitzpatrick, P. F. (1997). Identification of the naturally occurring flavin of nitroalkane oxidase from *Fusarium oxysporum* as a 5-nitrobutyl-FAD and conversion of the enzyme to the active FAD-containing form. *J. Biol. Chem., 272* (9), 5563-5570.

Gadda, G. and Fitzpatrick, P. F. (1998). Biochemical and physical characterization of the active FAD-containing form of nitroalkane oxidase from *Fusarium oxysporum*. *Biochemistry, 37* (17), 6154-6164.

Gadda, G. and Fitzpatrick, P. F. (1999). Substrate specificity of a nitroalkane oxidizing enzyme. *Arch. Biochem. Biophys., 363*, 309-313.

Gadda, G., Banerjee, A. and Fitzpatrick, P. F. (2000a). Identification of an essential tyrosine residue in nitroalkane oxidase by modification with tetranitromethane. *Biochemistry, 39* (5), 1162-1168.

Gadda, G., Banerjee, A., Dangott, L. J. and Fitzpatrick, P. F. (2000b). Identification of a cysteine residue in the active site of nitroalkane oxidase by modification with N-ethylmaleimide. *J. Biol. Chem., 275* (41), 31891-31895.

Gadda, G., Choe, D. Y. and Fitzpatrick, P. F. (2000d). Use of pH and kinetic isotope effects to dissect the effects of substrate size on binding and catalysis by nitroalkane oxidase. *Arch. Biochem. Biophys., 382* (1), 138-144.

Gadda, G. and Fitzpatrick, P. F. (2000e). Mechanism of nitroalkane oxidase: 2. pH and kinetic isotope effects. *Biochemistry, 39* (6), 1406-1410.

Gadda, G., Banerjee, A., Fleming, G. S. and Fitzpatrick, P. F. (2001). Evidence for an essential arginine in the flavoprotein nitroalkane oxidase. *J. Enzyme Inhib., 16* (2), 157-163.

Gatti, D. L., Palfey, B. A., Lah, M. S., Entsch, B., Massey, V., Ballou, D. P. and Ludwig, M. L. (1994). The mobile flavin of 4-OH benzoate hydroxylase. *Science, 266* (5182), 110-114.

Ghisla, S. and Massey, V. (1989). Mechanisms of flavoprotein-catalyzed reactions. *Eur. J. Biochem., 181* (1), 1-17.

Ghisla, S. and Thorpe, C. (2004). Acyl-CoA dehydrogenases. A mechanistic overview. *Eur. J. Biochem., 271* (3), 494-508.

Glykos, N. M. and Kokkinidis, M. (2001). Multidimensional molecular replacement. *Acta Cryst., D57*, 1462-1473.

Goodwin, A., Kersulyte, D., Sisson, G., Veldhuyzen van Zanten, S. J., Berg, D. E. and Hoffman, P. S. (1998). Metronidazole resistance in *Helicobacter pylori* is due to null mutations in a gene (rdxA) that encodes an oxygen-insensitive NADPH nitroreductase. *Mol. Microbiol., 28* (2), 383-393.

Gorlatova, N., Tchorzewski, M., Kurihara, T., Soda, K. and Esaki, N. (1998). Purification, characterization, and mechanism of a flavin mononucleotide-dependent 2-nitropropane dioxygenase from Neurospora crassa. *Appl. Environ. Microbiol., 64*, 1029-1033.

Gow, A. J., Farkouh, C. R., Munson, D. A., Posencheg, M. A. and Ischiropoulos, H. (2004). Biological significance of nitric oxide-mediated protein modifications. *Am. J. Physiol. Lung Cell Mol. Physiol., 287* (2), L262-268.

Guo, F., Chang, B. H. and Rizzo, C. J. (2002). An N1-hydrogen bonding model for flavin coenzyme. *Bioorg. Med. Chem. Lett., 12* (2), 151-154.

Haas, J. A. and Fox, B. G. (1999). Role of hydrophobic partitioning in substrate selectivity and turnover of the *Ricinus communis* stearoyl acyl carrier protein delta(9) desaturase. *Biochemistry, 38*, 12833-12840.

Hamilton, W. C., Rollett, J. S. and Sparks, R. A. (1965). On the relative scaling of X-ray photographs. *Acta Cryst., 18*, 129-130.

Hanawa, F., Tahara, S. and Towers, G. H. (2000). Antifungal nitro compounds from skunk cabbage (Lysichitum americanum) leaves treated with cupric chloride. *Phytochemistry, 53* (1), 55-58.

Hassel, B. and Sonnewald, U. (1995). Selective inhibition of the tricarboxylic acid cycle of GABAergic neurons with 3-nitropropionic acid in vivo. *J. Neurochem., 65* (3), 1184-1191.

Hauge, J. G., Crane, F. L. and Beinert, H. (1956). On the mechanism of dehydrogenation of fatty acyl derivatives of coenzyme A. III. Palmityl CoA dehydrogenase. *J. Biol. Chem., 219* (2), 727-733.

Hauptman, H. A. (1997). Shake-and-bake: an algorithm for automatic solution ab initio of crystal structures. *Methods Enzymol., 277*, 3-13.

Heasley, C. J. and Fitzpatrick, P. F. (1996). Kinetic mechanism and substrate specificity of nitroalkane oxidase. *Biochem. Biophys. Res. Commun., 225* (1), 6-10.

Heiss, G., Hofmann, K. W., Trachtmann, N., Walters, D. M., Rouviere, P. and Knackmuss, H. J. (2002). npd gene functions of *Rhodococcus (opacus) erythropolis* HL PM-1 in the initial steps of 2,4,6-trinitrophenol degradation. *Microbiology, 148* (Pt 3), 799-806.

Heiss, G., Trachtmann, N., Abe, Y., Takeo, M. and Knackmuss, H. J. (2003). Homologous npdGI genes in 2,4-dinitrophenol- and 4-nitrophenol-degrading *Rhodococcus* sp. *Appl. Environ. Microbiol., 69* (5), 2748-2754.

Hendrickson, W. A. and Ogata, C. M. (1997). Phase determination from multiwavelength anomalous diffraction measurements. *Methods Enzymol., 276*, 494-523.

Higgins, D., Thompson, J., Gibson, T., Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressivemultiple sequence alignment through sequence weighting,position-specific gap penalties and weight matrix choice. *Nucleic Acids Res., 22*, 4673-4680.

Hipkin, C. R., Salem, M. A., Simpson, D. and Wainwright, S. J. (1999). 3-nitropropionic acid oxidase from horseshoe vetch (*Hippocrepis comosa*): a novel plant enzyme. *Biochem. J., 340* (Pt 2), 491-495.

Homma-Takeda, S., Hiraku, Y., Ohkuma, Y., Oikawa, S., Murata, M., Ogawa, K., Iwamuro, T., Li, S., Sun, G. F., Kumagai, Y., Shimojo, N. and Kawanishi, S. (2002). 2,4,6-trinitrotoluene-induced reproductive toxicity via oxidative DNA damage by its metabolite. *Free Radic. Res., 36* (5), 555-566.

Howell, L. and Smith, D. (1992). Identification of heavy-atom derivatives by normal probability methods. *J. Appl. Cryst., 25*, 81-86.

Inestrosa, N. C., Bronfman, M. and Leighton, F. (1979). Properties of fatty acyl-CoA oxidase from rat liver, a peroxisomal flavoprotein. *Life Sci., 25* (13), 1127-1135.

Irie, Y., Saeki, M., Kamisaki, Y., Martin, E. and Murad, F. (2003). Histone H1.2 is a substrate for denitrase, an activity that reduces nitrotyrosine immunoreactivity in proteins. *Proc. Natl. Acad. Sci. U. S. A., 100* (10), 5634-5639.

Isildar, M., Abou-Khalil, W. H., Jimenez, J. J., Abou-Khalil, S. and Yunis, A. A. (1988). Aerobic nitroreduction of dehydrochloramphenicol by bone marrow. *Toxicol. Appl. Pharmacol., 94* (2), 305-310.

Jeong, J. Y., Mukhopadhyay, A. K., Dailidiene, D., Wang, Y., Velapatino, B., Gilman, R. H., Parkinson, A. J., Nair, G. B., Wong, B. C., Lam, S. K., Mistry, R., Segal, I., Yuan, Y.,

Gao, H., Alarcon, T., Brea, M. L., Ito, Y., Kersulyte, D., Lee, H. K., Gong, Y., Goodwin, A., Hoffman, P. S. and Berg, D. E. (2000). Sequential inactivation of rdxA (HP0954) and frxA (HP0642) nitroreductase genes causes moderate and high-level metronidazole resistance in *Helicobacter pylori*. *J. Bacteriol., 182* (18), 5082-5090.

Johnson, B. D. and Stankovich, M. T. (1993). Influence of two substrate analogues on thermodynamic properties of medium-chain acyl-CoA dehydrogenase. *Biochemistry, 32* (40), 10779-10785.

Johnson, G. R., Jain, R. K. and Spain, J. C. (2002). Origins of the 2,4-dinitrotoluene pathway. *J. Bacteriol., 184* (15), 4219-4232.

Jones, T. A., Zou, J. Y., Cowan, S. W. and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Cryst., A47*, 110-119.

Kabsch, W. and Sander, C. (1983). Dictionary of Protein Secondary Structure: Pattern Recognition of Hydrogen-Bonded and Geometrical Features. *Biopolymers, 22*, 2577-2637.

Kamisaki, Y., Wada, K., Bian, K., Balabanli, B., Davis, K., Martin, E., Behbod, F., Lee, Y. C. and Murad, F. (1998). An activity in rat tissues that modifies nitrotyrosine-containing proteins. *Proc. Natl. Acad. Sci. U. S. A., 95* (20), 11584-11589.

Kantardjieff, K. A. and Bernhard, R. (2003). Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein–nucleic acid complex crystals. *Prot. Sci., 12*, 1865-1871.

Karle, J. (1980). Some Developments in Anomalous Dispersion for the Structural Investigation of Macromolecular Systems in Biology. *International Journal of Quantum Chemistry: Quantum Biology Symposium, 7*, 357-367.

Karle, J. and Hauptman, H. A. (1950). The phases and magnitudes of the structure factors. *Acta Cryst., 3*, 181-187.

Khan, H., Barna, T., Harris, R. J., Bruce, N. C., Barsukov, I., Munro, A. W., Moody, P. C. and Scrutton, N. S. (2004). Atomic resolution structures and solution behaviour of enzyme-substrate complexes of *Enterobacter cloacae* PB2 pentaerythritol tetranitrate reductase: Multiple conformational states and implications for the mechanism of nitroaromatic explosive degradation. *J. Biol. Chem., 279* (29), 30563-30572.

Khan, H., Harris, R. J., Barna, T., Craig, D. H., Bruce, N. C., Munro, A. W., Moody, P. C. and Scrutton, N. S. (2002). Kinetic and structural basis of reactivity of pentaerythritol tetranitrate reductase with NADPH, 2-cyclohexenone, nitroesters, and nitroaromatic explosives. *J. Biol. Chem., 277* (24), 21906-21912.

Kido, T., Yamamoto, T. and Soda, K. (1975). Microbial assimilation of alkyl nitro compounds and formation of nitrite. *Arch. Microbiol., 106* (3), 165-169.

Kido, T., Soda, K., Suzuki, T. and Asada, K. (1976a). A new oxygenase, 2-nitropropane dioxygenase of *Hansenula mrakii*. Enzymologic and spectrophotometric properties. *J. Biol. Chem., 251* (22), 6994-7000.

Kido, T., Yamamoto, T. and Soda, K. (1976b). Purification and properties of nitroalkane-oxidizing enzyme from *Hansenula mrakii*. *J. Bacteriol., 126* (3), 1261-1265.

Kido, T., Soda, K. and Asada, J. (1978a). Properties of 2-nitropropane dioxygenase of Hansenula mrakii. Formation and participation of superoxide. *J. Biol. Chem., 253* (1), 226-232.

Kido, T., Hashizume, K. and Soda, K. (1978b). Purification and properties of nitroalkane oxidase from *Fusarium oxysporum*. *J. Bacteriol., 133* (1), 53-58.

Kido, T., Tanizawa, K., Inagaki, K., Yoshimura, T., Ishida, M., Hashizume, K. & Soda, K. (1984). 2-Nitropropane dioxygenase from Hansenula mrakii: re-characterization of the enzyme and oxidation of anionic nitroalkanes. *Agric. Biol. Chem., 48*, 2549–2554.

Kim, J. J. and Wu, J. (1988). Structure of the medium-chain acyl-CoA dehydrogenase from pig liver mitochondria at 3-A resolution. *Proc. Natl. Acad. Sci. U. S. A., 85* (18), 6677-6681.

Kim, J. J., Wang, M., Djordjevic, S. and Paschke, R. (1992). The three dimensional structure of acyl-CoA dehydrogenases. *Prog. Clin. Biol. Res., 375*, 111-126.

Kim, J. J., Wang, M. and Paschke, R. (1993). Crystal structures of medium-chain acyl-CoA dehydrogenase from pig liver mitochondria with and without substrate. *Proc. Natl. Acad. Sci. U. S. A., 90* (16), 7523-7527.

Kim, J. J. and Miura, R. (2004). Acyl-CoA dehydrogenases and acyl-CoA oxidases. Structural basis for mechanistic similarities and differences. *Eur. J. Biochem., 271* (3), 483-493.

Kirkpatrick, S., Gelatt Jr., C. D. and Vecchi, M. P. (1983). Optimization by Simulated Annealing. *Science, 220*, 671-680.

Kissinger, C., Gehlhaar, D. and Fogel, D. (1999). Rapid automated molecular replacement by evolutionary search. *Acta Cryst., D55*, 484-491.

Kleywegt, G. J. and Jones, T. A. (1998). Databases in protein crystallography. *Acta Cryst, D54*, 1119-1131.

Koeck, T., Fu, X., Hazen, S. L., Crabb, J. W., Stuehr, D. J. and Aulak, K. S. (2004). Rapid and selective oxygen-regulated protein tyrosine denitration and nitration in mitochondria. *J. Biol. Chem., 279* (26), 27257-27262.

Kohen, A. and Klinman, J. P. (1999). Hydrogen tunneling in biology. *Chem. Biol., 6* (7), 191 –198.

Krissinel, E. and Henrick, K. (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Cryst., D60* (Pt 12 Pt 1), 2256-2268.

Kuhn, D. M., Sakowski, S. A., Sadidi, M. and Geddes, T. J. (2004). Nitrotyrosine as a marker for peroxynitrite-induced neurotoxicity: the beginning or the end of the end of dopamine neurons? *J. Neurochem., 89* (3), 529-536.

Kunau, W. H., Dommes, V. and Schulz, H. (1995). β-oxidation of fatty acids in mitochondria, peroxisomes, and bacteria: a century of continued progress. *Prog. Lipid Res., 34* (4), 267-342.

Kurihara, T., Esaki, N., Soda, K. and Ohishi, N. (1994). A 340-nm chromophore of nitroalkane oxidase from *Fusarium oxysporum*. In K. Yagi (Ed.), *Flavins and flavoproteins* (pp. 195-198). Berlin, Germany: Walter de Gruyter.

Kurtz, K. A. and Fitzpatrick, P. F. (1997). pH and secondary kinetic isotope effects on the reaction of D-amino acid oxidase with nitroalkane anions: evidence for direct attack on the flavin by carbanions. *J. Am. Chem. Soc., 119* (5), 1155 –1156.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. and Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst., 26*, 283-291.

Lee, H. J., Wang, M., Paschke, R., Nandy, A., Ghisla, S. and Kim, J. J. (1996). Crystal structures of the wild type and the Glu376Gly/Thr255Glu mutant of human medium-chain acyl-CoA dehydrogenase: influence of the location of the catalytic base on substrate specificity. *Biochemistry, 35* (38), 12412-12420.

Lee, J. Y., Moon, S. S. and Hwang, B. K. (2003). Isolation and antifungal and antioomycete activities of aerugine produced by *Pseudomonas fluorescens* strain MM-B16. *Appl. Environ. Microbiol., 69* (4), 2023-2031.

Lewis, E. S. (1982). Nitro-activated carbon acids. In S. Patai (Ed.), *The chemistry of amino, nitroso and nitro compounds and their derivatives* (pp. 715-729). New York: John Wiley & Sons.

Little, H. N. (1951). Oxidation of nitroethane by extracts from *Neurospora*. *J. Biol. Chem., 193*, 347-358.

Lunin, V. Y., Lunina, N. L., Petrova, T. E., Skovoroda, T. P., Urzhumtsev, A. G. and Podjarny, A. D. (2000). Low-resolution ab initio phasing: problems and advances. *Acta Cryst., D56* (Pt 10), 1223-1232.

Lunin, V. Y., Lunina, N. L., Petrova, T. E., Urzhumtsev, A. G. and Podjarny, A. D. (1998). On the *ab initio* solution of the phase problem for macromolecules at very low resolution. II. Generalized Likelihood Based Approach to the Cluster Discrimination. *Acta Cryst., D54*, 726-734.

MacMicking, J., Xie, Q. W. and Nathan, C. (1997). Nitric oxide and macrophage function. *Annu. Rev. Immunol., 15*, 323-350.

Marconi, R. T., Lodmell, J. S. and Hill, W. E. (1990). Identification of a rRNA/chloramphenicol interaction site within the peptidyltransferase center of the 50 S subunit of the *Escherichia coli* ribosome. *J. Biol. Chem., 265* (14), 7894-7899.

Massey, V. (2000). The chemical and biological versatility of riboflavin. *Biochem. Soc. Trans., 28* (4), 283-296.

Matthews, B. W. (1968). Solvent content of protein crystals. *J. Mol. Biol., 33* (2), 491-497.

Meah, Y., Brown, B. J., Chakraborty, S. and Massey, V. (2001). Old yellow enzyme: reduction of nitrate esters, glycerin trinitrate, and propylene 1,2-dinitrate. *Proc. Natl. Acad. Sci. U. S. A., 98* (15), 8560-8565.

Mendz, G. L. and Megraud, F. (2002). Is the molecular basis of metronidazole resistance in microaerophilic organisms understood? *Trends Microbiol., 10* (8), 370-375.

Miyagi, M., Sakaguchi, H., Darrow, R. M., Yan, L., West, K. A., Aulak, K. S., Stuehr, D. J., Hollyfield, J. G., Organisciak, D. T. and Crabb, J. W. (2002). Evidence that light modulates protein nitration in rat retina. *Mol. Cell. Proteomics, 1* (4), 293-303.

Moran, G. R., Entsch, B., Palfey, B. A. and Ballou, D. P. (1997). Electrostatic effects on substrate activation in para-hydroxybenzoate hydroxylase: studies of the mutant lysine 297 methionine. *Biochemistry, 36* (24), 7548-7556.

Motamedi, H. and Shafiee, A. (1998). The biosynthetic gene cluster for the macrolactone ring of the immunosuppressant FK506. *Eur. J. Biochem., 256* (3), 528-534.

Muller, F. (1991). *In Chemistry and Biochemistry of Flavoenzymes* (Vol. 1). Boca Raton, FL: CRC Press.

Murshudov, G. N., Vagin, A. A. and Dodson, E. J. (1997). Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Cryst., D53*, 240-255.

Nagpal, A., Valley, M. P., Fitzpatrick, P. F. and Orville, A. M. (2004). Crystallization and preliminary analysis of active nitroalkane oxidase in three crystal forms. *Acta Cryst., D60*, 1456-1460.

Nagpal, A., Valley, M. P., Fitzpatrick, P. F. and Orville, A. M. (2005). Crystal Structures of Nitroalkane Oxidase Trapped During Turnover and an Inhibited Complex: Insights into the Structural Basis for Substrate Specificity and the Catalytic Mechanism. *Nat. Str. Mol. Biol., To be submitted*.

Nakajima, Y., Miyahara, I., Hirotsu, K., Nishina, Y., Shiga, K., Setoyama, C., Tamaoki, H. and Miura, R. (2002). Three-dimensional structure of the flavoenzyme acyl-CoA oxidase-II from rat liver, the peroxisomal counterpart of mitochondrial acyl-CoA dehydrogenase. *J. Biochem. (Tokyo), 131* (3), 365-374.

Nathan, C. and Shiloh, M. U. (2000). Reactive oxygen and nitrogen intermediates in the relationship between mammalian hosts and microbial pathogens. *Proc. Natl. Acad. Sci. U. S. A., 97* (16), 8841-8848.

Neuberger, G., Maurer-Stroh, S., Eisenhaber, B., Hartig, A. and Eisenhaber, F. (2003). Prediction of peroxisomal targeting signal 1 containing proteins from amino acid sequence. *J. Mol. Biol., 328* (3), 581-592.

Nivinskas, H., Koder, R. L., Anusevicius, Z., Sarlauskas, J., Miller, A. F. and Cenas, N. (2000). Two-electron reduction of nitroaromatic compounds by *Enterobacter cloacae* NAD(P)H nitroreductase: description of quantitative structure- activity relationships. *Acta Biochim. Pol., 47* (4), 941-949.

Ogata, C. M. (1998). MAD phasing grows up. *Nat. Struct. Biol., 5*, 638-640.

Ortiz-Maldonado, M., Cole, L. J., Dumas, S. M., Entsch, B. and Ballou, D. P. (2004). Increased positive electrostatic potential in p-hydroxybenzoate hydroxylase accelerates hydroxylation but slows turnover. *Biochemistry, 43* (6), 1569-1579.

Ortiz-Maldonado, M., Entsch, B. and Ballou, D. P. (2003). Conformational changes combined with charge-transfer interactions are essential for reduction in catalysis by p-hydroxybenzoate hydroxylase. *Biochemistry, 42* (38), 11234-11242.

Ortiz-Maldonado, M., Gatti, D., Ballou, D. P. and Massey, V. (1999). Structure-function correlations of the reaction of reduced nicotinamide analogues with p-hydroxybenzoate hydroxylase substituted with a series of 8-substituted flavins. *Biochemistry, 38* (50), 16636-16647.

Osato, T., Ueda, M., Fukuyama, S., Yagishita, K., Okami, Y. and Umezawa, H. (1955). Production of tertiomycin (a new antibiotic substance), azomycin and eurocidin by *S. eurocidicus*. *J. Antibiot. (Tokyo), 8* (4), 105-109.

Osumi, T., Hashimoto, T. and Ui, N. (1980). Purification and properties of acyl-CoA oxidase from rat liver. *J. Biochem. (Tokyo), 87* (6), 1735-1746.

Otwinowski, Z. and Minor, W. (1997). *Processing of X-ray Diffraction Data Collected in Oscillation Mode* (Vol. 276 Macromolecular Crystallography, part A). New York: Academic Press.

Pak, J. W., Knoke, K. L., Noguera, D. R., Fox, B. G. and Chambliss, G. H. (2000). Transformation of 2,4,6-trinitrotoluene by purified xenobiotic reductase B from *Pseudomonas fluorescens* I-C. *Appl. Environ. Microbiol., 66* (11), 4742-4750.

Palfey, B. A. and Massey, V. (1998). Flavin-dependent enzymes. In M. Sinnott (Ed.), *Comprehensive Biological Catalysis: A Mechanistic Reference* (Vol. III, pp. 83-154). San Diego: Academic Press.

Palfey, B. A., Moran, G. R., Entsch, B., Ballou, D. P. and Massey, V. (1999). Substrate recognition by "password" in p-hydroxybenzoate hydroxylase. *Biochemistry, 38* (4), 1153-1158.

Palmer, G. and Massey, V. (1968). In T. P. Singer (Ed.), *Biological Oxidation* (pp. 263-300). New York: John Wiley and Sons.

Parker, A. R. (2003). Binding of the human "electron transferring flavoprotein" (ETF) to the medium chain acyl-CoA dehydrogenase (MCAD) involves an arginine and histidine residue. *J. Enzyme. Inhib. Med. Chem., 18* (5), 453-462.

Pedersen, L. and Henriksen, A. (2005). Acyl-CoA oxidase 1 from *Arabidopsis thaliana*. Structure of a key enzyme in plant lipid metabolism. *J. Mol. Biol., 345* (3), 487-500.

Pellett, J. D., Becker, D. F., Saenger, A. K., Fuchs, J. A. and Stankovich, M. T. (2001). Role of aromatic stacking interactions in the modulation of the two-electron reduction potentials of flavin and substrate/product in Megasphaera elsdenii short-chain acyl-coenzyme A dehydrogenase. *Biochemistry, 40* (25), 7720-7728.

Perrakis, A., Morris, R. M. and Lamzin, V. S. (1999). Automated protein model building combined with iterative structure refinement. *Nat. Struct. Biol., 6*, 458-463.

Podjarny, A. D. and Urzhumtsev, A. (1997). *Low resolution phasing* (Vol. 276): Academic Press.

Porter, D. J. and Bright, H. J. (1980). 3-Carbanionic substrate analogues bind very tightly to fumarase and aspartase. *J. Biol. Chem., 255* (10), 4772-4780.

Porter, D. J. and Bright, H. J. (1987). Propionate-3-nitronate oxidase from *Penicillium atrovenetum* is a flavoprotein which initiates the autoxidation of its substrate by O_2. *J. Biol. Chem., 262* (30), 14428-14434.

Porter, D. J., Voet, J. G. and Bright, H. J. (1973). Direct evidence for carbanions and covalent N 5 -flavin-carbanion adducts as catalytic intermediates in the oxidation of nitroethane by D-amino acid oxidase. *J. Biol. Chem., 248* (12), 4400-4416.

Powell, H. R. (1999). The Rossmann Fourier autoindexing algorithm in *MOSFLM*. *Acta Cryst., D55* (Pt 10), 1690-1695.

Ramakrishanan, C. and Ramachandran, G. N. (1965). Stereochemical criteria for polypeptide and protein chain conformations, Part II, Allowed conformations for a pair of peptide units. *Biophy. J., 5* (909-933).

Randy, R. J. (2001). Pushing the boundaries of molecular replacement with maximum likelihood. *Acta Cryst., D57* (Pt 10), 1373-1382.

Read, R. J. (1986). Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Cryst., A42*, 140-149.

Reibenspies, J. H., Guo, F. and Rizzo, C. J. (2000). X-ray crystal structures of conformationally biased flavin models. *Org. Lett., 2* (7), 903-906.

Rice, L. M. and Brünger, A. T. (1994). Torsion angle dynamics: reduced variable conformational sampling enhances crystallographic structure refinement. *Proteins, 19*, 277-290.

Rieger, P. G., Meier, H. M., Gerle, M., Vogt, U., Groth, T. and Knackmuss, H. J. (2002). Xenobiotics in the environment: present and future strategies to obviate the problem of biological persistence. *J. Biotechnol., 94* (1), 101-123.

Rizzo, C. J. (2001). Further computational studies on the conformation of 1,5-dihydrolumiflavin. *Antioxid. Redox Signal, 3* (5), 737-746.

Roberts, D. L., Frerman, F. E. and Kim, J. J. (1996). Three-dimensional structure of human electron transfer flavoprotein to 2.1 Å resolution. *Proc. Natl. Acad. Sci. U. S. A., 93* (25), 14355-14360.

Roberts, P., Basran, J., Wilson, E. K., Hille, R. and Scrutton, N. S. (1999). Redox cycles in trimethylamine dehydrogenase and mechanism of substrate inhibition. *Biochemistry, 38* (45), 14927-15940.

Rosser, S. J., Basran, A., Travis, E. R., French, C. E. and Bruce, N. C. (2001). Microbial transformations of explosives. *Adv. Appl. Microbiol., 49*, 1-35.

Rossmann, M. G. (2001). Molecular replacement-historical background. *Acta Cryst., D57* (Pt 10), 1360-1366.

Rossmann, M. G. and Blow, D. M. (1962). The detection of subunits within the crystallographic asymmetric unit. *Acta Cryst., 15*, 24-31.

Sampson, N. S. (2001). Dissection of a Flavoenzyme Active Site: The Reaction Catalyzed by Cholesterol Oxidase. *Antioxidants & Redox Signaling, 3* (5), 839-846.

Samuelson, J. (1999). Why metronidazole is active against both bacteria and parasites. *Antimicrob. Agents Chemother., 43* (7), 1533-1541.

Satoh, A., Nakajima, Y., Miyahara, I., Hirotsu, K., Tanaka, T., Nishina, Y., Shiga, K., Tamaoki, H., Setoyama, C. and Miura, R. (2003). Structure of the transition state analog of medium-chain acyl-CoA dehydrogenase. Crystallographic and molecular orbital studies on the charge-transfer complex of medium-chain acyl-CoA dehydrogenase with 3-thiaoctanoyl-CoA. *J. Biochem. (Tokyo), 134* (2), 297-304.

Schomaker, V. and Trueblood, K. N. (1968). On the rigid-body motion of molecules in crystals. *Acta Cryst., B24*, 63-76.

Schopfer, F. J., Baker, P. R. and Freeman, B. A. (2003). NO-dependent protein nitration: a cell signaling event or an oxidative inflammatory response? *Trends Biochem. Sci., 28* (12), 646-654.

Schreuder, H. A., van der Laan, J. M., Hol, W. G. and Drenth, J. (1988). Crystal structure of *p*-hydroxybenzoate hydroxylase complexed with its reaction product 3,4-dihydroxybenzoate. *J. Mol. Biol., 199* (4), 637-648.

Schreuder, H. A., Prick, P. A., Wierenga, R. K., Vriend, G., Wilson, K. S., Hol, W. G. and Drenth, J. (1989). Crystal structure of the p-hydroxybenzoate hydroxylase-substrate complex refined at 1.9 Å resolution. Analysis of the enzyme-substrate and enzyme-product complexes. *J. Mol. Biol., 208* (4), 679-696.

Schreuder, H. A., van der Laan, J. M., Swarte, M. B., Kalk, K. H., Hol, W. G. and Drenth, J. (1992). Crystal structure of the reduced form of p-hydroxybenzoate hydroxylase refined at 2.3 A resolution. *Proteins, 14* (2), 178-190.

Schreuder, H. A., Mattevi, A., Obmolova, G., Kalk, K. H., Hol, W. G., van der Bolt, F. J. and van Berkel, W. J. (1994). Crystal structures of wild-type p-hydroxybenzoate hydroxylase complexed with 4-aminobenzoate,2,4-dihydroxybenzoate, and 2-hydroxy-4-aminobenzoate and of the Tyr222Ala mutant complexed with 2-hydroxy-4-aminobenzoate. Evidence for a proton channel and a new binding mode of the flavin ring. *Biochemistry, 33* (33), 10161-10170.

Schuettelkopf, A. W. and van Aalten, D. M. F. (2004). PRODRG - a tool for high-throughput crystallography of protein-ligand complexes. *Acta Cryst., D60*, 1355--1363.

Setoyama, C., Tamaoki, H., Nishina, Y., Shiga, K. and Miura, R. (1995). Functional expression of two forms of rat acyl-CoA oxidase and their substrate specificities. *Biochem. Biophys. Res. Commun., 217* (2), 482-487.

Sheldrick, G. M. and Gould, R. O. (1995). Structure solution by iterative peaklist optimization and tangent expansion in space group P1. *Acta Cryst., B51*, 423-431.

Sheriff, S., Klei, H. and Davis, M. (1999). Implementation of a six-dimensional search using the amore translation function for difficult molecular replacement problems. *J. Appl. Cryst., 32*, 98-101.

Shishehbor, M. H. and Hazen, S. L. (2004). Inflammatory and oxidative markers in atherosclerosis: relationship to outcome. *Curr. Atheroscler. Rep., 6* (3), 243-250.

Shoji, J. H., Hinoo, H., Terui, Y., Kikuchi, J., Hattori, T., Ishii, K., Matsumoto, K. and Yoshida, T. (1989). Isolation of azomycin from *Pseudomonas fluorescens*. *J. Antibiot. (Tokyo), 42* (10), 1513-1514.

Sinnokrot, M. O. and Sherrill, C. D. (2004). Substituent Effects in Pi-Pi Interactions: Sandwich and T-shaped Configurations. *J. Am. Chem. Soc., 126*, 7690-7697.

Sisson, G., Goodwin, A., Raudonikiene, A., Hughes, N. J., Mukhopadhyay, A. K., Berg, D. E. and Hoffman, P. S. (2002). Enzymes associated with reductive activation and action of nitazoxanide, nitrofurans, and metronidazole in *Helicobacter pylori*. *Antimicrob. Agents Chemother., 46* (7), 2116-2123.

Sisson, G., Jeong, J. Y., Goodwin, A., Bryden, L., Rossler, N., Lim-Morrison, S., Raudonikiene, A., Berg, D. E. and Hoffman, P. S. (2000). Metronidazole activation is mutagenic and causes DNA fragmentation in *Helicobacter pylori* and in *Escherichia coli* containing a cloned *H. pylori* RdxA(+) (Nitroreductase) gene. *J. Bacteriol., 182* (18), 5091-5096.

Somerville, C. C., Nishino, S. F. and Spain, J. C. (1995). Purification and characterization of nitrobenzene nitroreductase from *Pseudomonas pseudoalcaligenes* JS45. *J. Bacteriol., 177* (13), 3837-3842.

Spain, J. C. (1995). Biodegradation of nitroaromatic compounds. *Annu. Rev. Microbiol., 49*, 523-555.

Spain, J. C., Hughes, J. B. and Knackmuss, H. J. (2000). *Biodegradation of Nitroaromatic Compounds and Explosives*. Boca Raton: CRC Press LLC.

Spiess, T., Desiere, F., Fischer, P., Spain, J. C., Knackmuss, H. J. and Lenke, H. (1998). A new 4-nitrotoluene degradation pathway in a *Mycobacterium* strain. *Appl. Environ. Microbiol., 64* (2), 446-452.

Stankovich, M. T. (1991). *In Chemistry and Biochemistry of Flavoenzymes* (Vol. 1). Boca Raton, FL: CRC Press.

Stankovich, M. T. and Fox, B. G. (1984). Redox potential-pH properties of the flavoprotein L-amino-acid oxidase. *Biochim. Biophys. Acta, 786* (1-2), 49-56.

Stankovich, M. T., Sabaj, K. M. and Tonge, P. J. (1999). Structure/function of medium chain acyl-CoA dehydrogenase: the importance of substrate polarization. *Arch. Biochem. Biophys., 370* (1), 16-21.

Steyn-Parve, E. P. and Beinert, H. (1958a). On the mechanism of dehydrogenation of fatty acyl derivatives of coenzyme A. VI. Isolation and properties of stable enzyme-substrate complexes. *J. Biol. Chem., 233* (4), 843-852.

Sussman, J. L., Holbrook, S. R., Church, G. M. and Kim, S. H. (1977). Structure-factor least-squares refinement procedure for macromolecular structure using constrained and restrained parameters. *Acta Cryst., A33*, 800-804.

Suzuki, K., Gomi, T., Kaidoh, T. and Itagaki, E. (1991). Hydroxylation of *o*-halogenophenol and *o*-nitrophenol by salicylate hydroxylase. *J. Biochem. (Tokyo), 109* (2), 348-353.

Swoboda, B. E. and Massey, V. (1965). Purification and Properties of the glucose oxidase from *Asperillus niger*. *J. Biol. Chem., 240*, 2209-2215.

Tanaka, K. and Indo, Y. (1992). Evolution of the acyl-CoA dehydrogenase/oxidase superfamily. *Prog. Clin. Biol. Res., 375*, 95-110.

Terwilliger, T. C. and Eisenberg, D. (1983). Unbiased three-dimensional refinement of heavy-atom parameters by correlation of origin-removed Patterson functions. *Acta Cryst., A39*, 813-817.

Terwilliger, T. C., Kim, S. H. and Eisenberg, D. (1987). Generalized method of determining heavy-atom positions using the difference Patterson function. *Acta Cryst., A43*, 1-5.

Terwilliger, T. C. and Berendzen, J. (1997). Bayesian Correlated MAD phasing. *Acta Cryst., D53*, 571-579.

Terwilliger, T. C. and Berendzen, J. (1999). Automated MAD and MIR structure solution. *Acta Cryst., D55* (Pt 4), 849-861.

Terwilliger, T. C. (2001). Maximum-likelihood density modification using pattern recognition of structural motifs. *Acta Cryst., D57*, 1755-1762.

Terwilliger, T. C. (2002). Automated structure solution, density modification and model building. *Acta Cryst., D58*, 1937-1940.

Terwilliger, T. C. (2004). Using prime-and-switch phasing to reduce model bias in molecular replacement. *Acta Cryst., D60*.

Thorpe, C. and Kim, J. P. (1995). Structure and mechanism of action of the Acyl-CoA dehydrogenases. *FASEB J., 9*, 718-725.

Tiffany, K. A., Roberts, D. L., Wang, M., Paschke, R., Mohsen, A. W., Vockley, J. and Kim, J. J. (1997). Structure of human isovaleryl-CoA dehydrogenase at 2.6 Å resolution: structural basis for substrate specificity. *Biochemistry, 36* (28), 8455-8464.

Tong, L. and Rossmann, M. G. (1997). Rotation function calculations with GLRF program. *Methods Enzymol., 276*, 594-611.

Toogood, H. S., Van Thiel, A., Basran, J., Sutcliffe, M. J., Scrutton, N. S. and Leys, D. (2004). Extensive domain motion and electron transfer in the human ETF-MCAD complex. *J. Biol. Chem., 279* (31), 32904-32912.

Trinh, S. and Reysset, G. (1998). Mutagenic action of 5-nitroimidazoles: in vivo induction of GC-->CG transversion in two *Bacteroides fragilis* reporter genes. *Mutat. Res., 398* (1-2), 55-65.

Trott, S., Nishino, S. F., Hawari, J. and Spain, J. C. (2003). Biodegradation of the Nitramine Explosive CL-20. *Appl. Environ. Microbiol., 69* (3), 1871-1874.

Tunez, I., Munoz, M. C. and Montilla, P. (2005). Treatment with Dehydroepiandrosterone Prevents Oxidative Stress Induced by 3-Nitropropionic Acid in Synaptosomes. *Pharmacology, 74* (3), 113-118.

Turko, I. V., Li, L., Aulak, K. S., Stuehr, D. J., Chang, J. Y. and Murad, F. (2003). Protein tyrosine nitration in the mitochondria from diabetic mouse heart. Implications to dysfunctional mitochondria in diabetes. *J. Biol. Chem., 278* (36), 33972-33977.

Turko, I. V. and Murad, F. (2002). Protein nitration in cardiovascular diseases. *Pharmacol. Rev., 54* (4), 619-634.

Urzhumtsev, A. G., Lunina, N. L., Skovoroda, T. P., Podjarny, A. D. and Lunin, V. Y. (2000). Density constraints and low-resolution phasing. *Acta Cryst., D56* (Pt 10), 1233-1244.

Urzhumtsev, A. G., Vernoslova, E. A. and Podjarny, A. D. (1996). Approaches to very low resolution phasing of the ribosome 50S particle form Thermus thermophilus by the Few-Atoms-Models and Molecular-Replacement Methods. *Acta Cryst., D52*, 1092-1097.

Vagin, A. and Teplyakov, A. (1997). MOLREP: an automated program for molecular replacement. *J. Appl. Cryst., 30*, 1022-1025.

Valley, M. P. and Fitzpatrick, P. F. (2003a). Inactivation of nitroalkane oxidase upon mutation of the active site base and rescue with a deprotonated substrate. *J. Am. Chem. Soc., 125* (29), 8738-8739.

Valley, M. P. and Fitzpatrick, P. F. (2003b). Reductive half-reaction of nitroalkane oxidase: effect of mutation of the active site aspartate to glutamate. *Biochemistry, 42* (19), 5850-5856.

Valley, M. P. and Fitzpatrick, P. F. (2004). Comparison of enzymatic and non-enzymatic nitroethane anion formation: thermodynamics and contribution of tunneling. *J. Am. Chem. Soc., 126* (20), 6244-6245.

Valley, M. P., Tichy, S. E. and Fitzpatrick, P. F. (2005). Establishing the kinetic competency of the cationic imine intermediate in nitroalkane oxidase. *J. Am. Chem. Soc., 127* (7), 2062-2066.

van der Vliet, A., Eiserich, J. P., Kaur, H., Cross, C. E. and Halliwell, B. (1996). Nitrotyrosine as biomarker for reactive nitrogen species. *Methods Enzymol., 269*, 175-184.

van Pee, K. H. (1996). Biosynthesis of halogenated metabolites by bacteria. *Annu. Rev. Microbiol., 50*, 375-399.

Wallace, A. C., Laskowski, R. A. and Thornton, J. M. (1995). LIGPLOT: A program to generate schematic diagrams of protein-ligand interactions. *Prot. Eng., 8*, 127-134.

Walsh, C. T. (2004). Polyketide and nonribosomal peptide antibiotics: modularity and versatility. *Science, 303* (5665), 1805-1810.

Wang, B. C. (1985). Resolution of the Phase Ambiguity in Macromolecular Crystallography. *Methods Enzymol., 115*, 90-112.

Watanabe, K., Khosla, C., Stroud, R. M. and Tsai, S. C. (2003). Crystal structure of an Acyl-ACP dehydrogenase from the FK520 polyketide biosynthetic pathway: insights into extender unit biosynthesis. *J. Mol. Biol., 334* (3), 435-444.

Weeks, C. M., DeTitta, G. T., Miller, R. and Hauptman, H. A. (1993). Application of the minimal principle to peptide structures. *Acta Cryst., D49* (Pt 1), 179-181.

Wierenga, R. K., de Jong, R. J., Kalk, K. H., Hol, W. G. and Drenth, J. (1979). Crystal structure of *p*-hydroxybenzoate hydroxylase. *J. Mol. Biol., 131* (1), 55-73.

Williams, R. E., Rathbone, D. A., Moody, P. C., Scrutton, N. S. and Bruce, N. C. (2001). Degradation of explosives by nitrate ester reductases. *Biochem. Soc. Symp.* (68), 143-153.

Williams, R. E. and Bruce, N. C. (2000). The role of nitrate ester reductase enzymes in the biodegradation of explosives. In J. C. Spain, J. B. Hughes, andH. J. Knackmuss (Eds.), *Biodegradation of nitroaromatic compounds and explosives* (pp. 161-184). Boca Raton: Lewis Publishers.

Williams, R. E. and Bruce, N. C. (2002). 'New uses for an Old Enzyme'--the Old Yellow Enzyme family of flavoenzymes. *Microbiology, 148* (Pt 6), 1607-1614.

Williams, R. E., Rathbone, D. A., Scrutton, N. S. and Bruce, N. C. (2004). Biotransformation of explosives by the old yellow enzyme family of flavoproteins. *Appl. Environ. Microbiol., 70* (6), 3566-3574.

Winn, M. D., Isupov, M. N. and Murshudov, G. N. (2001). Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Cryst., D57*, 122-133.

Wolfenden, R., Snider, M., Ridgway, C. and Miller, B. (1999). The temperature dependence of enzyme rate enhancements. *J. Am. Chem. Soc., 121*, 7419 –7420.

Wu, J., Bell, A. F., Luo, L., Stephens, A. W., Stankovich, M. T. and Tonge, P. J. (2003). Probing hydrogen-bonding interactions in the active site of medium-chain acyl-CoA dehydrogenase using Raman spectroscopy. *Biochemistry, 42* (40), 11846-11856.

Wu, K., Chung, L., Revill, W. P., Katz, L. and Reeves, C. D. (2000). The FK520 gene cluster of *Streptomyces hygroscopicus var. ascomyceticus* (ATCC 14891) contains genes for biosynthesis of unusual polyketide extender units. *Gene, 251* (1), 81-90.

Xu, W. M. and Liu, L. Z. (1998). Nitric oxide: from a mysterious labile factor to the molecule of the Nobel Prize. Recent progress in nitric oxide research. *Cell Res., 8* (4), 251-258.

Ye, Y. Z., Strong, M., Huang, Z. Q. and Beckman, J. S. (1996). Antibodies that recognize nitrotyrosine. *Methods Enzymol., 269*, 201-209.

Yin, Y., Sampson, N. S., Vrielink, A. and Lario, P. I. (2001). The presence of a hydrogen bond between asparagine 485 and the pi system of FAD modulates the redox potential in the reaction catalyzed by cholesterol oxidase. *Biochemistry, 40* (46), 13779-13787.

Zeelen, J. P., Hiltunen, J. K., Ceska, T. A. and Wierenga, R. K. (1994). Crystallization experiments with 2-enoyl-CoA hydratase, using an automated 'fast-screening' crystallization protocol. *Acta Cryst., D50*, 443-447.

Zhang, J. and Tan, H. (2002). Cloning, expression and characterization of a gene encoding nitroalkane- oxidizing enzyme from *Streptomyces ansochromogenes*. *Eur. J. Biochem., 269* (24), 6302-6307.

Zhang, X. and Matthews, B. W. (1995). EdPDB: A Multi-Functional Tool for Protein Structure Analysis. *J. Appl. Cryst., 28*, 624-630.

VITA

Akanksha Nagpal was born on December 20, 1975 in Delhi, India, to Harish Nagpal and Rekha Nagpal. After schooling in St. Anthony's Senior Secondary School (Delhi), Dayanand Anglo Vedic School (Delhi) and J.D. Tytler School (Delhi), she obtained her Bachelor of Science in Biology in the year 1995 from Hans Raj College, University of Delhi, India. She attended Indian Institute of Technology, Roorkee, to pursue a M.S. in Biotechnology (1997). After finishing her master's degree, she worked on Hindustan Lever Ltd. sponsored projects as a Research Scientist from 1997–2000. She came to Georgia Tech in the year 2000 to pursue a Ph.D., specializing in protein crystallography, under the supervision of Dr. Allen M. Orville. To realize her entrepreneurial dreams and to hone her business acumen, she concurrently pursued business management as a minor. While at Georgia Tech, she co-founded a start up venture, ProFold, and served as its CEO.